Evolutionary versus instrumental goals: How evolutionary psychology misconceives human rationality

Keith E. Stanovich  
University of Toronto, Canada

Richard F. West  
James Madison University, Harrisonburg, Virginia, USA

An important research tradition in the cognitive psychology of reasoning—called the heuristics and biases approach—has firmly established that people’s responses often deviate from the performance considered normative on many reasoning tasks. For example, people assess probabilities incorrectly, they display confirmation bias, they test hypotheses inefficiently, they violate the axioms of utility theory, they do not properly calibrate degrees of belief, they overproject their own opinions onto others, they display illogical framing effects, they uneconomically honour sunk costs, they allow prior knowledge to become implicated in deductive reasoning, and they display numerous other information processing biases (for summaries of the large literature, see Baron, 1998, 2000; Dawes, 1998; Evans, 1989; Evans & Over, 1996; Kahneman & Tversky, 1972, 1984, 2000; Kahneman, Slovic, & Tversky, 1982; Nickerson, 1998; Shafir & Tversky, 1995; Stanovich, 1999; Tversky, 1996).

It has been common for these empirical demonstrations of a gap between descriptive and normative models of reasoning and decision making to be taken as indications that systematic irrationalities characterize human cognition. However, over the last decade, an alternative interpretation of these findings has been championed by various evolutionary psychologists, adaptationist modellers, and ecological theorists (Anderson, 1990, 1991; Chater & Oaksford, 2000; Cosmides & Tooby, 1992, 1994a, 1996; Gigerenzer, 1996a; Oaksford & Chater, 1998, 2001; Rode, Cosmides, Heck, & Tooby, 1999; Todd & Gigerenzer, 2000). They have reinterpreted the modal response in most

of the classic heuristics and biases experiments as indicating an optimal information processing adaptation on the part of the subjects. It is argued by these investigators that the research in the heuristics and biases tradition has not demonstrated human irrationality at all and that a Panglossian position (see Stanovich & West, 2000), which assumes perfect human rationality, is the proper default position to take.

It will be argued in this chapter that although the work of the evolutionary psychologists has uncovered some fundamentally important things about human cognition, these theorists have misconstrued the nature of human rationality and have conflated important distinctions in this domain. What these theorists have missed (or failed to sufficiently emphasize) is that definitions of rationality must coincide with the level of the entity whose optimization is at issue. This admonition plays out most directly in the distinction between evolutionary rationality and instrumental rationality—necessitated by the fact that the optimization procedures for replicators and for vehicles (to use Dawkins' (1976) terms) need not always coincide. The distinction follows from the fact that the genes—as subpersonal replicators—can increase their fecundity and longevity in ways that do not necessarily serve the instrumental goals of the vehicles built by the genome (Skyrms, 1996; Stanovich, 1999). Despite their frequent acknowledgements that the conditions in the environment of evolutionary adaptedness (EEA) do not match those of modern society, evolutionary psychologists have a tendency to background potential mismatches between genetic interests and personal interests.

We will argue below that dual-process models of cognitive functioning provide a way of reconciling the positions of the evolutionary psychologists and researchers in the heuristics and biases tradition. Such models acknowledge the domain specificity of certain modular processes emphasized by the evolutionary psychologists. But importantly, they also posit general, interactive, nonautonomous, and central serial-processing operations of executive control and problem solving that serve to guarantee instrumental rationality by overriding the responses generated by autonomous modules when the latter threaten optimal outcomes at the personal level.

DEBATES ABOUT THE NORMATIVE RESPONSE IN HEURISTICS AND BIASES TASKS: SOME EXAMPLES

The empirical data pattern that provoked our attempted reconciliation of the positions of the evolutionary psychologists and researchers in the heuristics and biases tradition is the repeated finding in our research (Stanovich & West, 1998a, 1998b, 1998c, 1998d, 1999, 2000) that the modal response was different from the response given by the more cognitively able subjects. We have related this finding to the disputes about which response is normative in various heuristics and biases investigated task in the entire Wason's (1966) selection task. table; two show letters and two each card has a number on one experimenter has the following respect to the four cards: "If there is an even number on the they must turn over whichever experimenter's rule is true or of the selection task is extreme Manktelow, 1999; Newstead & of participants make the cor (not-Q)—the only two cards t incorrect choices made by p (P and Q) or the selection of Ql

Numerous alternative expla and P responses have been gi Johnson-Laird, 1999; Liberman 1995; Oaksford & Cl Evans, 1995; Oaksford & Cl Stanovich & West, 1998a). W several of these alternative t results from the operation of a example, Oaksford and Chate that rather than interpreting ( the experimenter intends), many probabilistic hypothesis testing P and Q response is actually th the problem is assumed along al. (1995) stress that selection tive mechanisms. They expl ential comprehension mechanism optimally relevant communica

Our second example of the heuristics and biases research by the much-investigated Linc

Linda is 31 years old, single, o sophy. As a student, she was di social justice, and also parti the following statements by 1 and 8 for the least probable.
ents as indicating an optimal of the subjects. It is argued by ristics and biases tradition has nd that a Panglossian position effect human rationality, is the l the work of the evolutionary tally important things about nstrued the nature of human inctions in this domain. What ally emphasize) is that defi-level of the entity whose opti-most directly in the distinction ental rationality—necessitated for replicators and for vehicles ways coincide. The distinction subpersonal replicators—can s that do not necessarily serve the genome (Skyrms, 1996; owledgements that the condi-pittedness (EEA) do not match ogists have a tendency to back-interests and personal interests. dels of cognitive functioning the evolutionary psychologists adition. Such models acknow- processes emphasized by the they also posit general, inter-crossing operations of executive instrumental rationality itonomous modules when the all level.

ACTIVE RESPONSE SES TASKS:

ES attempted reconciliation of the id researchers in the heuristics ur research (Stanovich & West, the modal response was differ- nitively able subjects. We have high response is normative in various heuristics and biases tasks. An example is provided by the most investigated task in the entire reasoning and problem solving literature— Wason’s (1966) selection task. The participant is shown four cards laying on a table; two show letters and two show numbers (A, D, 3, 8). They are told that each card has a number on one side and a letter on the other and that the experimenter has the following rule (of the "if P, then Q" type) in mind with respect to the four cards: "If there is a vowel on one side of the card, then there is an even number on the other side". The participants are then told that they must turn over whichever cards are necessary to determine whether the experimenter's rule is true or false. Performance on such abstract versions of the selection task is extremely low (Evans, Newstead, & Byrne, 1993; Manktelow, 1999; Newstead & Evans, 1995). Typically, less than 10 per cent of participants make the correct selections of the A card (P) and 3 card (not-Q)—the only two cards that could falsify the rule. The most common incorrect choices made by participants are the A card and the 8 card (P and Q) or the selection of the A card only (P).

Numerous alternative explanations for the preponderance of incorrect PQ and P responses have been given (see Evans et al., 1993; Hardman, 1998; Johnson-Laird, 1999; Liberman & Klar, 1996; Margolis, 1987; Newstead & Evans, 1995; Oaksford & Chater, 1994; Sperber, Cara, & Girotto, 1995; Stanovich & West, 1998a). What is important in the present context is that several of these alternative theories posit that the incorrect PQ response results from the operation of efficient and optimal cognitive mechanisms. For example, Oaksford and Chater (1994, 1996; see also Nickerson, 1996) argue that rather than interpreting the task as one of deductive reasoning (as the experimenter intends), many people interpret it as an inductive problem of probabilistic hypothesis testing (see Evans & Over, 1996). They show that the P and Q response is actually the expected one if an inductive interpretation of the problem is assumed along with optimal data selection (which they modeled with a Bayesian analysis). Although their model is different, Sperber et al. (1995) stress that selection task performance is driven by optimized cognitive mechanisms. They explain selection task performance in terms of inferential comprehension mechanisms that are "geared towards the processing of optimally relevant communicative behaviors" (Sperber et al., 1995, p. 90).

Our second example of theorists defending as rational the response that heuristics and biases researchers have long considered incorrect is provided by the much-investigated Linda Problem (Tversky & Kahneman, 1983):

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.
(a) Linda is a teacher in an elementary school
(b) Linda works in a bookstore and takes Yoga classes
(c) Linda is active in the feminist movement
(d) Linda is a psychiatric social worker
(e) Linda is a member of the League of Women Voters
(f) Linda is a bank teller
(g) Linda is an insurance salesperson
(h) Linda is a bank teller and is active in the feminist movement

Because alternative (h) (Linda is a bank teller and is active in the feminist movement) is the conjunction of alternatives (c) and (f), the probability of (h) cannot be higher than that of either (c) (Linda is active in the feminist movement) or (f) (Linda is a bank teller), yet 85 per cent of the participants in Tversky and Kahneman’s (1983) study rated alternative (h) as more probable than (f), thus displaying the so-called conjunction fallacy. Those investigators argued that logical reasoning on the problem (all feminist bank tellers are also bank tellers, so (h) cannot be more probable than (f)) was trumped by a heuristic based on so-called representativeness that primes answers to problems based on an assessment of similarity (a feminist bank teller seems to overlap more with the description of Linda than does the alternative “bank teller”). Of course, logic dictates that the subset (feminist bank teller/superset (bank teller)) relationship should trump assessments of representativeness when judgements of probability are at issue.

However, several investigators have suggested that rather than illogical cognition, it is rational pragmatic inferences that lead to the violation of the logic of probability theory in the Linda Problem (see Adler, 1991; Dunlap & Hilton, 1991; Politzer & Noveck, 1991; Slugoski & Wilson, 1998). Hilton (1995, p. 264) summarizes the view articulated in these critiques by arguing that “the inductive nature of conversational inference suggests that many of the experimental results that have been attributed to faulty reasoning may be reinterpreted as being due to rational interpretations of experimenter-given information”.

In short, these critiques imply that displaying the conjunction fallacy is a rational response triggered by the adaptive use of social cues, linguistic cues, and background knowledge (see Hilton, 1995). For example, Macdonald and Gilbody (1990, p. 59) argue that it is possible that subjects will:

...usually assume the questioner is asking the question because there is some reason to suppose that Linda might be a bank teller and the questioner is interested to find out if she is... If Linda were chosen at random from the electoral register and “bank teller” was chosen at random from some list of occupations, the probability of them corresponding would be very small, certainly less than 1 in 100...the question itself has suggested to the subjects that Linda could be a feminist bank teller. Subjects are therefore being asked to judge how likely it is that Linda is a feminist bank teller when there is some

unknown reason to suppose so itself.

Hilton (1995; see Dunlap & Hi subjects’ behaviour on the Lin detailed information given ab knows a considerable amount a the phrase “Linda is a bank te active in the feminist moveme this to be the case. If “Linda is rating (h) as more probable t fallacy.

Several investigators have shown seeming violations of the logic (see Adler, 1984, 1991; Hertwig, Slugoski & Wilson, 1998). Mos (1975) norms of rational comm ber & Wilson, 1986; Sperber et cooperative with the listener— attempt to be cooperative is by ing the so-called Grican max understand a speaker’s meanin meaning of what is spoken bu assuming that the speaker inte at pains to remind us that th cognition. They are rational he cuts as emphasized in the heu not to be seen as processing more efficient processing modes pp. 265–266).

However, it is not clear why in should cause respondents to at correct and rational... Incenti conversationally rational interp context... the conversational increased incentives lead respor rational in the context.

Clearly, in the view of these the such contexts does not represent

Many theorists have linked formance to the automatic lin theorists commonly posit that th
unknown reason to suppose she is, which reason has prompted the question itself.

Hilton (1995; see Dulaney & Hilton, 1991) provides a similar explanation of subjects' behaviour on the Linda Problem. Under the assumption that the detailed information given about the target means that the experimenter knows a considerable amount about Linda, then it is reasonable to think that the phrase "Linda is a bank teller" does not contain the phrase "and is not active in the feminist movement" because the experimenter already knows this to be the case. If "Linda is a bank teller" is interpreted in this way, then rating (h) as more probable than (f) no longer represents a conjunction fallacy.

Several investigators have suggested that pragmatic inferences lead to seeming violations of the logic of probability theory in the Linda Problem (see Adler, 1984, 1991; Hertwig & Gigerenzer, 1999; Politzer & Noveck, 1991; Slokoski & Wilson, 1998). Most of these can be analysed in terms of Grice's (1975) norms of rational communication (see Hilton & Slokoski, 2000; Sperber & Wilson, 1986; Sperber et al., 1995), which require that the speaker be cooperative with the listener— and one of the primary ways that speakers attempt to be cooperative is by not being redundant. The key to understanding the so-called Gricean maxims of communication is to realize that to understand a speaker's meaning the listener must comprehend not only the meaning of what is spoken but also what is implicated in a given context assuming that the speaker intends to be cooperative. And Hilton (1995) is at pains to remind us that these are rational aspects of communicative cognition. They are rational heuristics as opposed to the suboptimal shortcuts as emphasized in the heuristics and biases literature. Thus, they are not to be seen as processing modes that are likely to be given up for more efficient processing modes when the stakes become high (Hilton, 1995, pp. 265-266);

However, it is not clear why increasing the financial stakes in an experiment should cause respondents to abandon an interpretation that is pragmatically correct and rational. Incentives are not going to make respondents drop a conversationally rational interpretation in favor of a less plausible one in the context . . . the conversational inference approach does not predict that increased incentives lead respondents to change an interpretation that seems rational in the context.

Clearly, in the view of these theorists, committing the conjunction fallacy in such contexts does not represent a cognitive error.

Many theorists have linked their explanation of Linda-problem performance to the automatic linguistic socialization of information. These theorists commonly posit that the socialization tendency reflects evolutionary
adaptations in the domain of social intelligence. This linkage stems from many theories that, although varied in their details, all posit that much of human intelligence has foundations in social interaction (Baldwin, 2000; Barton & Dunbar, 1997; Brothers, 1990; Bugental, 2000; Byrne & Whiten, 1988; Caporael, 1997; Cosmides, 1989; Cummins, 1996; Dunbar, 1998; Humphrey, 1976; Jolly, 1966; Kummer, Daston, Gigerenzer, & Silk, 1997; Mithen, 1996; Tomasello, 1999; Whiten & Byrne, 1997).

In a seminal essay that set the stage for this hypothesis, Nicholas Humphrey (1976) argued that the impetus for the development of primate intelligence was the need to master the social world. Based on his observation of nonhuman primates, Humphrey (1976) concluded that the knowledge and information processing necessary to engage efficiently with the physical world seemed modest compared with the rapidly changing demands of the social world. Humphrey (1976) posited that the latter was the key aspect of the environment that began to bootstrap higher intelligence in all primates.

This social, or interactional, intelligence forms that substrate upon which all future evolutionary and cultural developments in modes of thought are overlaid. That such social intelligence forms the basic substrate upon which all higher forms of intelligence must build leads to the important assumption that a social orientation toward problems is always available as a default processing mode when computational demands become onerous. The cognitive illusions demonstrated by three decades of work in problem solving, reasoning, and decision making (Evans, 1989; Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 1996, 2000; Stanovich, 1999) seem to bear this out. As in the Linda Problem and four-card selection task discussed above, the literature is full of problems where an abstract, decontextualized—but computationally expensive—approach is required for the normatively appropriate answer. However, often, alongside such a solution, resides a tempting social approach ("Oh, yeah, the author of this knows a lot about Linda") that with little computational effort will prime a response.

Since our theme has now been established with the selection-task and Linda-problem examples, our final two examples of theorists defending as rational the response that heuristics and biases researchers have long considered incorrect will be described only briefly.

Covariation detection

The 2 \times 2 covariation detection task is run in a variety of different formats (Levin, Wasserman, & Kao, 1993; Stanovich & West, 1998d; Wasserman, Dorner, & Kao, 1990). In one, for example, subjects are asked to evaluate the efficacy of a drug based on a hypothetical well-designed scientific experiment. They are told that:

These data correspond to four ally labelled A, B, C, and D (evaluate the effectiveness of to detect that the drug is ineffectively harmful. Only 50 per cured (150 out of 300), but 80 percent cure rate (300 out of 375).

Much previous experimental jects weight the cell information: cell D receiving the least weight. 1983; Kao & Wasserman, 1993: to ignore cell D is non-normal weight the four cells. The norr Kao & Wasserman, 1993: Shapurel—subtracting from the p indicator is present the probabi- tor is absent. Numerically, the problem presented above is as follows:

Despite the fact that it is in such experiments underweig (Stanovich & West, 1998d) cell the 2 \times 2 contingency assess adapted information process in- onstrates that an adaptive mix the D cell (cause absent and of Green, 2001) and containing no discrepancy weighting of the cells in a 2 \times 2 Thus, here again in another ti nonnormative—but that resp an adaptationist analysis.

Probability matching

The probabilistic contingency (Gal & Baron, 1996; Tversky front of two lights (one red an
gence. This linkage stems from details, all posit that much of ial interaction (Baldwin, 2000; igentual, 2000; Byrne & Whiten, ummins, 1996; Dunbar, 1998; ton, Gigerenzer, & Silk, 1997; rne, 1997). This hypothesis, Nicholas Hum e development of primate intel- lworld. Based on his observation concluded that the knowledge nage efficiently with the phy- ices rapidly changing demands of d that the latter was the key strap higher intelligence in all orts that substrate upon which ments in modes of thought are the basic substrate upon which ds to the important assumption s always available as a default ds become onerous. The cogni- s of work in problem solving, ; Kahneman, Slovic, & Tversky, novich, 1999) seem to bear this selection task discussed above, abstract, decontextualized—but required for the normatively ide such a solution, resides a thor of this knows a lot about ill prime a response. ed with the selection-task and ames of theorists defending l biases researchers have long ilefully.

In a variety of different formats h & West, 1998d; Wasserman, objects are asked to evaluate the -designed scientific experiment.

7. EVOLUTIONARY VERSUS INSTRUMENTAL GOALS

150 people received the drug and were cured.
150 people received the drug and were not cured.
300 people did not receive the drug and were cured.
75 people did not receive the drug and were not cured.

These data correspond to four cells of the $2 \times 2$ contingency table traditionally labelled A, B, C, and D (see Levin et al., 1993). Subjects are asked to evaluate the effectiveness of the drug on a scale. In this case, they have to detect that the drug is ineffective. In fact, not only is it ineffective, it is positively harmful. Only 50 per cent of the people who received the drug were cured (150 out of 300), but 80 per cent of those who did not receive the drug were cured (300 out of 375).

Much previous experimentation has produced results indicating that subjects weight the cell information in the order cell A > cell B > cell C > cell D—cell D receiving the least weight and/or attention (see Arkes & Harkness, 1983; Kao & Wasserman, 1993; Schustack & Sternberg, 1981). The tendency to ignore cell D is non-normative, as indeed is any tendency to differentially weight the four cells. The normatively appropriate strategy (see Allen, 1980; Kao & Wasserman, 1993; Shanks, 1995) is to use the conditional probability rule—subtracting from the probability of the target hypothesis when the indicator is present the probability of the target hypothesis when the indicator is absent. Numerically, the rule amounts to calculating the $\Delta p$ statistic: $[A/(A + B)] - [C/(C + D)]$ (see Allen, 1980). For example, the $\Delta p$ value for the problem presented above is $-0.30$, indicating a fairly negative association.

Despite the fact that it is a nonnormative strategy, the modal subject in such experiments underweights (sometimes markedly underweights, see Stanovich & West, 1998d) cell D. However, Anderson (1990) has modelled the $2 \times 2$ contingency assessment experiment using a model of optimally adapted information processing and come to a startling conclusion. He demonstrates that an adaptive model can predict the much-replicated finding that the D cell (cause absent and effect absent) is vastly underweighted (but see Over & Green, 2001) and concludes that “this result makes the point that there need be no discrepancy between a rational analysis and differential weighting of the cells in a $2 \times 2$ contingency table” (Anderson, 1990, p. 160). Thus, here again in another task is the pattern where the modal response is nonnormative—but that response has been defended from the standpoint of an adaptationist analysis.

Probability matching

The probabilistic contingency experiment has many versions in psychology (Gal & Baron, 1996; Tversky & Edwards, 1966). In one, the subjects sit in front of two lights (one red and one blue) and are told that they are to predict
which of the lights will be flashed on each trial and that there will be several
dozens of such trials (subjects are often paid money for correct predictions).
The experimenter has actually programmed the lights to flash randomly, with
the provision that the red light will flash 70 per cent of the time and the blue
light 30 per cent of the time. Subjects do quickly pick up the fact that the red
light is flashing more, and they predict that it will flash on more trials than
they predict that the blue light will flash. Most often, they switch back and
forth, predicting the red light roughly 70 per cent of the time and the blue
light roughly 30 per cent of the time.

This strategy of probability matching is suboptimal because it insures that,in
this example, the subject will correctly predict only 58 per cent of the time
 \((0.7 \times 0.7 + 0.3 \times 0.3)\) compared with the 70 per cent hit rate achieved by predicting
the more likely colour on each trial. In fact, much experimentation has indicated that animals and humans often fail to maximize expected utility in the probabilistic contingency experiment¹ (Estes, 1964, 1976; Gallistel, 1990;
Tversky & Edwards, 1966). Nevertheless, Gigerenzer (1996b; see also Cooper,
1989) shows how probability matching could, under some conditions, actual-
ly be an evolutionarily stable strategy (see Skyrms, 1996, for many such
examples). Thus, we have in probability matching our final example of how
a non-normative response tendency is defended on an evolutionary or
adaptationist account.

**DISSOCIATIONS BETWEEN COGNITIVE ABILITY AND THE MODAL RESPONSE IN HEURISTICS AND BIASES TASKS**

We will argue in this chapter that, in each of these examples, evolutionary rationality has dissociated from normative rationality—where the latter is
viewed as utility maximization for the individual organism (instrumental
rationality) and the former is defined as survival probability at the level of
the gene (Dawkins, 1976, 1982). Our conceptualization of these findings
explicitly acknowledges the impressive record of descriptive accuracy enjoyed
by a variety of adaptationist and evolutionary models in predicting the modal
response (Anderson, 1990, 1991; Gigerenzer, 1996b; Oaksford & Chater,
1994, 1996; Rode et al., 1999), but our account attempts to make sense of another important empirical fact—that cognitive ability often dissoc-
iates from the response deemed adaptive on an evolutionary analysis

¹ In probability learning, or choice situations, both animals and humans generally approxi-
mate probability matching when reinforcement is delivered on variable-interval schedules.
However, things appear to be more complex when reinforcement is delivered on variable-ratio
schedules. Although humans still tend to approximate probability matching (Estes, 1964, 1976,
1984), animals often maximize (Herrnstein & Loveland, 1975; MacDonell, 1988, but see
Gallistel; 1990; Graf, Bullock, & Boterman, 1964; Sutherland & Mackintosh, 1971).
NATIVE ABILITY VS. HEURISTICS

These examples, evolutionary saliency—where the latter is sal, organism (instrumental probability at the level of the internalization of these findings) descriptive accuracy enjoyed / models in predicting the enzer, 1996b; Oaksford & Chater attempt to make cognitive ability often dispositional analysis and humans generally approximated on variable-interval schedules is delivered on variable-ratio variability matching (Estes, 1964, 1976, 1975; MacDonald, 1988; see 1 & MacIntosh, 1971). (Stanovich, 1999; Stanovich & West, 2000). Specifically, we have repeatedly found that in cases where the normative response is not the modal response, the subjects in the sample who were the highest in cognitive ability gave the normative response rather than the modal response. This is true for each of the four tasks described above.

For example, Table 7.1 presents the results from an investigation of ours (Stanovich & West, 1998a) using a selection task with a nonequivalent rule, the so-called Destination rule (in this instance: If "Baltimore" is on one side of the ticket, then "plane" is on the other side of the ticket). The table presents the mean SAT scores for several of the dominant choices on this selection rule (the SAT test is a test used for university admissions in the United States that is highly loaded on psychometric g). From the table, it is clear that respondents giving the deductively correct P and not-Q response had the highest SAT scores, followed by the subjects choosing the P card only. All other responses, including the modal P and Q response (chosen by 49 per cent of the sample), were given by subjects having SAT scores almost 100 points lower than those giving the correct response under a deductive construal. It is to the credit of models of optimal data selection (Oaksford & Chater, 1994) that they predict the modal response. But we are left with the seemingly puzzling finding that the response deemed optimal under such an analysis (PQ) is given by subjects of substantially lower general intelligence than the minority giving the response deemed correct under a strictly deductive interpretation of the problem (PNQ).

A similar puzzle surrounds findings on the Linda conjunction problem. Gricean analyses assume that those subjects committing the conjunction fallacy in such a contrived problem are reflecting the evolved use of socio-linguistic cues. Because this group is in fact the vast majority in most studies—and because the use of such pragmatic cues and background knowledge is often interpreted as reflecting adaptive information processing (Hilton, 1995)—it might be expected that these individuals would be the subjects of higher cognitive ability. We found the contrary. In our study
(Stanovich & West, 1998b), we examined the performance of 150 subjects on the Linda Problem. Consistent with the results of previous experiments on this problem (Tversky & Kahneman, 1983), 80.7 per cent of our sample committed the conjunction effect—they rated the feminist bank teller alternative as more probable than the bank teller alternative. However, the mean SAT score of the 121 subjects who committed the conjunction fallacy was 82 points lower than the mean score of the 29 who avoided the fallacy. This difference was highly significant and it translated into an effect size of .746 (which Rosenthal & Rosnow, 1991, classify as large). Thus, the pragmatic interpretations of why the conjunction effect is the modal response on this task might well be correct, but the modal response happens not to be the one given by the most intelligent subjects in the sample.

Likewise, in the $2 \times 2$ covariation detection experiment, we have found (Stanovich & West, 1998d) that it is those subjects weighting cell D more equally (not those underweighting the cell in the way that the adaptationist model dictates) who are higher in cognitive ability and who tend to respond normatively on other tasks. Again, Anderson (1990, 1991) might well be correct that a rational model of information processing in the task predicts underweighting of cell D by most subjects, but more severe underweighting is in fact associated with lower cognitive ability in our individual differences analyses.

Finally, we found a similar pattern in several experiments on probability matching using a variety of different paradigms (West & Stanovich, in press). For example, in one experiment involving choices among general strategies for approaching the probabilistic prediction task, subjects were given the following task description:

A die with 4 red faces and 2 green faces will be rolled 60 times. Before each roll you will be asked to predict which color (red or green) will show up once the die is rolled. You will be given one dollar for each correct prediction. Assume that you want to make as much money as possible. What strategy would you use in order to make as much money as possible by making the most correct predictions?

They were asked to choose from among the following five strategies:

Strategy A: Go by intuition, switching when there has been too many of one color or the other.
Strategy B: Predict the more likely color (red) on most of the rolls but occasionally, after a long run of reds, predict a green.
Strategy C: Make predictions according to the frequency of occurrence (4 of 6 for red and 2 of 6 for green). That is, predict twice as many reds as greens.
Strategy D: Predict the more likely color (red) on all of the 60 rolls.

7. EVOLUTIVE STRATEGIES

Strategy E: Predict more red depending upon 

The probability matching strategy that is optimally optimal strategy is presented in Table 7.2. The strategies were both preferred and their mean SAT scores. The strategies were both preferred and the modal choice. Again, the probability matching strategy is the maximizing strategy, as before, it is the maximizing strategy of the subjects with the highest probability choosing the maximizing preferred probability matching

RECONCILING THE TWO

We see in the results just reviewed. The evolutionary psychology correctly predicts the modal red. Yet in all of these cases—despite the modal response quite strate associations that also m the nondeontic selection task (and Q) are higher in cognitive cung the opposite response, proportionately tend to adher formation experiment, it is who are higher in cognitive a

<table>
<thead>
<tr>
<th>Mean SAT total on a probabilistic</th>
<th>Strategy choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>suit</td>
<td>Strategy A</td>
</tr>
<tr>
<td>suit</td>
<td>Strategy B</td>
</tr>
<tr>
<td>suit</td>
<td>Strategy C**</td>
</tr>
<tr>
<td>suit</td>
<td>Strategy D</td>
</tr>
<tr>
<td>suit</td>
<td>Strategy E</td>
</tr>
<tr>
<td>* = The probability correctly suit</td>
<td>** = The probability correctly suit</td>
</tr>
</tbody>
</table>


The performance of 150 subjects on a test of previous strategies on a 10.7 per cent of our sample of the feminist bank teller alternative. However, the mean of the conjunction fallacy was 82 who avoided the fallacy. This led to an effect size of .746 as large. Thus, the pragmatic is the modal response on this one happens not to be the one to one typical experiment, we have found subjects weighing cell D more the way that the adaptationist utility and who tend to respond on (1990, 1991) might well be processing in the task predicts more severe underweighting is one in our individual differences in, e.g., our general prediction task, subjects were

rolled 60 times. Before each roll (green) will show up once the die (correct prediction. Assume that what strategy would you use in by making the most correct

allowing five strategies:

1. Most of the rolls but dis, predict a green.
2. Frequency of occurrence (4 of it is, predict twice as many reds on all of the 60 rolls.

Strategy E: Predict more red than green, but switching back and forth depending upon "runs" of one color or the other.

The probability matching strategy corresponds to Strategy C here, and the normatively optimal strategy is Strategy D, which maximizes expected utility. Table 7.2 presents the number of subjects choosing each of the five strategies and their mean SAT scores. The probability matching and maximizing strategies were both preferred over the three foil strategies, with the former being the modal choice. Again, the choice defensible on evolutionary grounds (probability matching, see Gigerenzer, 1996b), is the modal choice. But again, as before, it is the maximizing, normatively dictated choice that is the choice of the subjects with the highest intellectual ability. The mean SAT scores of those choosing the maximizing choice was 55 points higher than those who preferred probability matching (p < .001).

RECONCILING THE TWO DATA PATTERNS WITHIN A TWO-PROCESS VIEW

We see in the results just reviewed two basic patterns that must be reconciled. The evolutionary psychologists and optimal data selection theorists correctly predict the modal response in a host of heuristics and biases tasks. Yet in all of these cases—despite the fact that the adaptationist models predict the modal response quite well—individual differences analyses demonstrate associations that also must be accounted for. Correct responders on the non-deontic selection task (P and not-Q choosers—not those choosing P and Q) are higher in cognitive ability. Despite conversational implicatures cuing the opposite response, individuals of higher cognitive ability disproportionately tend to adhere to the conjunction rule. In the 2 x 2 covariation detection experiment, it is those subjects weighting cell D more equally who are higher in cognitive ability. Finally, subjects of higher intelligence

<table>
<thead>
<tr>
<th>Strategy choice</th>
<th>SAT score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy A</td>
<td>1151 (15)</td>
</tr>
<tr>
<td>Strategy B</td>
<td>1163 (64)</td>
</tr>
<tr>
<td>Strategy C*</td>
<td>1160 (168)</td>
</tr>
<tr>
<td>Strategy D**</td>
<td>1215 (150)</td>
</tr>
<tr>
<td>Strategy E</td>
<td>1148 (48)</td>
</tr>
</tbody>
</table>

* = The probability matching response; ** = The normatively correct utility maximizing response.
disproportionally avoid the evolutionarily justified probability matching tendency.

We believe that a useful framework for incorporating both of these data patterns is provided by two-process theories of reasoning (Epstein, 1994; Evans, 1984, 1986; Evans & Over, 1996; Sloman, 1996; Stanovich, 1999). Such a framework can encompass both the impressive record of descriptive accuracy enjoyed by a variety of evolutionary/adaptationist models as well as the fact that cognitive ability sometimes dissociates from the response deemed optimal on an adaptationist analysis.

A summary of terms used by several two-process theorists and the generic properties distinguished by several two-process views is presented in Table 7.3. Although the details and technical properties of these dual-process theories do not always match exactly, nevertheless there are clear family resemblances (for discussion, see Evans & Over, 1996; Gigerenzer & Regier, 1996; Sloman, 1996). To emphasize that his concept of these two processes involved the synthesis of a prototype of the different models in the literature (rather than to attempt to defend the specific and unique properties of any one view), Stanovich (1999) adopted the generic labels System 1 and System 2.

The key differences in the properties of the two systems are listed in Table 7.3. System 1 processes are characterized as automatic, heuristic-based, and relatively undemanding of computational capacity. Thus, System 1 processes conjoin properties of automaticity, modularity, and heuristic processing as these constructs have been variously discussed in the literature. There is a sense in which the term System 1 is a misnomer in that it implies that it is referring to a single system. In fact, we intend the term System 1 to refer to a (probably large) set of systems in the brain (partially encapsulated modules in some views) that operate autonomously—in response to their own triggering stimuli and not under the control of a central processing structure (System 2).

System 2 coexists with the various characteristics that have been viewed as typifying controlled processing—serial, rule-based, language-biased, computationally expensive cognition. System 2 encompasses the processes of analytic intelligence that have traditionally been studied by information processing theorists trying to uncover the computational components underlying intelligence. Evans and Over (1999) argue that the function of the explicit processes of System 2 is to support hypothetical thinking. In their view, hypothetical thinking involves representing possible states of the world rather than actual states of affairs; for example, “deductive reasoning is hypothetical when its premises are not actual beliefs, but rather assumptions or suppositions….” Consequential decision making consists of forecasting a number of possible future world states and representing the possible actions available… Scientific thinking is itself hypothetical when entertaining hypotheses about the way the world might be and deducing their consequences for making predictions” (Evans & Over, 1999, p. 764). Evans and
TABLE 7.3
The terms for the two systems used by a variety of theorists and the properties of dual-process theories of reason

<table>
<thead>
<tr>
<th>Dual-process theories</th>
<th>System 1 (TASS)</th>
<th>System 2 (Analytic system)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sloman (1996)</td>
<td>Associative system</td>
<td>Rule-based system</td>
</tr>
<tr>
<td>Evans &amp; Over (1996)</td>
<td>Tacit thought processes</td>
<td>Explicit thought processes</td>
</tr>
<tr>
<td>Evans (1984, 1989)</td>
<td>Heuristic processing</td>
<td>Analytical processing</td>
</tr>
<tr>
<td>Evans &amp; Wason (1976)</td>
<td>Type 1 processes</td>
<td>Type 2 processes</td>
</tr>
<tr>
<td>Reber (1993)</td>
<td>Implicit cognition</td>
<td>Explicit learning</td>
</tr>
<tr>
<td>Levinson (1995)</td>
<td>Interational intelligence</td>
<td>Analytic intelligence</td>
</tr>
<tr>
<td>Epstein (1994)</td>
<td>Experiential system</td>
<td>Rational system</td>
</tr>
<tr>
<td>Pollock (1991)</td>
<td>Quick and inflexible modules</td>
<td>Intuition</td>
</tr>
<tr>
<td>Klein (1998)</td>
<td>Recognition-printed decisions</td>
<td>Rational choice strategy</td>
</tr>
<tr>
<td>Johnson-Laird (1983)</td>
<td>Implicit inferences</td>
<td>Explicit inferences</td>
</tr>
<tr>
<td>Fodor (1983)</td>
<td>Modular processes</td>
<td>Central processes</td>
</tr>
<tr>
<td>Chaiten, Liberman, &amp;</td>
<td>Heuristic processing</td>
<td>Systematic processing</td>
</tr>
<tr>
<td>Eagly (1983)</td>
<td>Animal control system</td>
<td></td>
</tr>
<tr>
<td>Gibbard (1990)</td>
<td>Contention scheduling</td>
<td></td>
</tr>
<tr>
<td>Norman &amp; Shallice</td>
<td>Automatic processing</td>
<td></td>
</tr>
<tr>
<td>(1986)</td>
<td>Automatic activation</td>
<td></td>
</tr>
<tr>
<td>Shiffrin &amp; Schneider</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1977)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posner &amp; Snyder (1975)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Properties

- Sloman: Associative, Holistic, Automatic, Relatively undemanding of cognitive capacity, Relatively fast, Acquisition by biology, exposure, and personal experience, Highly contextualized, Short-leash genetic goals that are relatively stable
- Evans: Rule-based, Analytic, Controlled, Demanding of cognitive capacity, Relatively slow, Acquisition by cultural and formal tuition, Decontextualized, Long-leash goals that are utility-maximizing for the organism and constantly updated because of changes in environment

TASS = The Autonomous Set of Systems.

Over (1999) posit that hypothetical thought involves representing assumptions—and these necessarily must be represented as such; otherwise content would be conformed with belief. Linguistic forms such as conditionals provide a medium for such representations—and the serial manipulation of this type of representation seems to be largely a System 2 function. Language provides the discrete representational tools that fully exploit the
computational power of the serial manipulations of which System 2 is capable (following Dennett, 1991, we think that System 2 processing is computationally demanding because the serial processes must be simulated by a largely parallel network).

The two systems tend to lead to different types of task construals. Construals triggered by System 1 are highly contextualized, personalized, and socialized. They are driven by considerations of relevance and are aimed at inferring intentionality by the use of conversational implicature even in situations that are devoid of conversational features (see Hilton, 1995). These properties characterize what Levinson (1995) has termed intentional intelligence—a system composed of the mechanisms that support a Greccan theory of communication that relies on intention-attribution. The primacy of these mechanisms leads to what has been termed the fundamental computational bias in human cognition (Stanovich, 1999, in press)—the tendency or predilection toward automatic contextualization of problems. In contrast, the more controlled processes of System 2 serve to decontextualize and depersonalize problems. This system is more adept at representing in terms of rules and underlying principles. It can deal with problems without social content and is not dominated by the goal of attributing intentionality nor by the search for conversational relevance.

Using the distinction between System 1 and System 2 processing, Stanovich and West (2000) argued that in order to observe large cognitive ability differences in a problem situation, the two systems must strongly cue different responses. One reason that this outcome is predicted is that it is assumed that individual differences in System 1 processes (interactional intelligence) are smaller and bear little relation to individual differences in System 2 processes (analytic intelligence; see McGeorge, Crawford, & Kelly, 1997; Reber, 1993; Reber, Walkenfeld, & Herstadt, 1991). If the two systems cue opposite responses, rule-based System 2 will tend to differentially cue those of high analytic intelligence and this tendency will not be diluted by System 1 nondifferentially drawing subjects to the same response. For example, in noneconomic selection tasks there is ample opportunity for the two systems to cue different responses. A deductive interpretation conjoined with an exhaustive search for 3 We strongly caution that the term "bias" is used throughout this chapter to denote "a preponderating disposition or propensity" (The Compact Edition of the Oxford Short English Dictionary, p. 211) and not a processing error. That a processing bias does not necessarily imply a cognitive error is a point repeatedly emphasized by the critics of the heuristics and biases literature (Funder, 1987; Gigerenzer, 1996a; Hastie & Raaijkes, 1988; Kruglanski & Ajzen, 1983), but in fact it was always the position of the original heuristics and biases researchers themselves (Kahneman, 2000; Kahneman & Tversky, 1973, 1996; Tversky & Kahneman, 1974). Thus, the use of the term bias here is meant to connote "default value" rather than "error". Under the assumption that computational biases result from evolutionary adaptations of the brain (Cosmides & Tooby, 1994b), it is likely that they are efficacious in most situations.

falsifying instances yields the result that processing style is likely associated within the heuristic-analytic framework response of P and Q of Evans' theory, a linguistically cue.

The sampling of experimental for further examples indicates the critics of the heuristics and b jects of lower analytic intelligence alternative construals may have the intuitional sense—as the evolution higher in a more flexible type of tively engaged, see Smith & Lev prone to follow normative rules 7.

EPILOGUE

The argument depends on the concept of instrumental rationality (at least in Dawkins' (1976, 1982) terms, every process relevant to the so-called mental rationality concerns utility of a task, to use Hull's, 1982, (1990, 1991) emphases this d model in psychology: Anders Skyrms, 1996) argument that evolutionary rationality) does not instrumental sense that is focus result, a descriptive model of pr deviate substantially from a n (Skyrms, 1996, spends an entire may be different models charac personal levels, respectively. A key aspect of our framework 1 It should be noted that the distinct different from the distinction between (1996). They define rationality, as means efficient for achieving one's goals" (p. 8
ions of which System 2 is hat System 2 processing is processes must be simulated pes of task construals. Con- onal implicature even in situ- ational implicature and relevance and are aimed at ure at decontextualization and at representing in terms of the problems without social judging intentionality nor by System 2 processing. Stano-

ferse large cognitive abilities is must strongly cue different; this is that it is assumed that interactional intelligence) are rences in System 2 processes and (Kelly, 1997; Reber, 1993; two systems cue opposite in different ways; that is, high dilution by System 1 is non-differentials. For example, in non-semantic two systems to cue different with an exhaustive search for throughout this chapter to denote "a list of the Oxford Short English bias does not necessarily imply any of the heuristics and biases litera-
ture; Kunglowski & Ajzen, 1983; but and biases researchers themselves rely & Kahneman, 1974). Thus, the phrase "not" rather than "error." Under these conditions, the brain is in most situations.

falsifying instances yields the response P and not-Q. This interpretation and processing style is likely associated with the rule-based System 2. In contrast, within the heuristic-analytic framework of Evans (1984, 1989, 1996), the matching response of P and Q reflects the heuristic processing of System 1 (in Evans' theory, a linguistically cued relevance response).

The sampling of experimental results reviewed here (see Stanovich, 1999, for further examples) indicates that the alternative responses favoured by the critics of the heuristics and biases literature were the choices of the subjects of lower analytic intelligence. We will explore the possibility that these alternative construals may have been triggered by heuristics that make evolution-ary sense—as the evolutionary psychologists argue—but that subjects higher in a more flexible type of analytic intelligence (and those more cognitively engaged, see Smith & Levin, 1996, Stanovich & West, 1999) are more prone to follow normative rules that maximize personal utility.

**EVOLUTIONARY RATIONALITY IS NOT INSTRUMENTAL RATIONALITY**

The argument depends on the distinction between evolutionary adaptation and instrumental rationality (utility maximization given goals and beliefs). The key point is that for the latter (variously termed practical, pragmatic, or means/ends rationality), maximization is at the level of the individual person. Adaptive optimization in the former case is at the level of the genes. In Dawkins' (1976, 1982) terms, evolutionary adaptation concerns optimization processes relevant to the so-called replicators (the genes), whereas instrumental rationality concerns utility maximization for the so-called vehicle (or interactors, to use Hull's, 1982, term), which houses the genes. Anderson (1990, 1991) emphasizes this distinction in his treatment of adaptationist models in psychology. Anderson (1990) accepts Stich's (1990; see also Skyrms, 1996) argument that evolutionary adaptation (hereafter termed evolutionary rationality) does not guarantee perfect human rationality in the instrumental sense that is focused on goals of the whole organism. As a result, a descriptive model of processing that is adaptively optimal could well deviate substantially from a normative model of instrumental rationality (Skyrms, 1996, spends an entire book demonstrating just this) because there may be different models characterizing optimization at the subpersonal and personal levels, respectively.1

A key aspect of our framework is the assumption that the goal structures

---

1 It should be noted that the distinction between evolutionary and instrumental rationality is different from the distinction between rationality, and rationality, utilized by Evans and Over (1996). They define rationality, as reasoning and acting "in a way that is generally reliable and efficient for achieving one's goals" (p. 8). Rationality, concerns reasoning and acting "when one
that are keyed to primarily the genes’ interests and the goal structures keyed primarily to the organism’s interests are differentially represented in Systems 1 and 2 (see Reber, 1992, 1993, for a theoretical and empirical basis for this claim). It is hypothesized that the features of System 1 are designed to very closely track increases in the reproduction probability of genes. System 2, while also clearly an evolutionary product, is primarily a control system focused on the interests of the whole person. It is the primary maximizer of an individual’s personal utility. Maximizing the latter will occasionally result in sacrificing genetic fitness (Barkow, 1989; Cooper, 1989; Skyrms, 1996). Because System 2 is more attuned to instrumental rationality than is System 1, System 2 will seek to fulfill the individual’s goals in the minority of cases where those goals conflict with the responses triggered by System 1.

Thus, the last difference between Systems 1 and 2 listed in Table 7.3 is that System 1 instantiates short-leash genetic goals, whereas System 2 instantiates a flexible goal hierarchy that is oriented toward maximizing goal satisfaction at the level of the whole organism. We borrow the short- /long-leash terminology by way of another metaphor used by Dawkins (1976), Dennett (1984), and Plotkin (1988)—the “Mars Explorer” analogy. Dennett (1984) describes how, when controlling a device such as a model airplane, one’s sphere of control is only limited by the power of the equipment, but when the distances become large, the speed of light becomes a non-negligible factor. NASA engineers responsible for the Mars Explorer vehicle knew that direct control was impossible because “the time required for a round trip signal was greater than the time available for appropriate action . . . Since controllers on Earth could no longer reach out and control them, they had to control themselves” (Dennett, 1984, p. 55, italics in the original). The NASA engineers had to move from the “short-leash” direct control, as in the model airplane case, to the “long-leash” control of the Mars Explorer case where the vehicle is not given moment-by-moment instructions on how to act, but instead is given a more flexible type of intelligence plus some generic goals.

As Dawkins (1976), in his similar discussion of the Mars Explorer logic in the science fiction story A for Andromeda notes, there is an analogy here to the type of control exerted by the genes when they build a brain: “The genes can only do their best in ad-hoc themselves . . . Like the chess survival machines not in spec the living trade . . . The ad- greatly cuts down on the nun the original program” (Dawkins according to Dawkins (1976), trend towards the emancip makers from their ultimate n vival machines and their ne power over behavior. But the next are taken by the nervou brains are the executives reached in any species, would single overall policy instructi The type of long-leash c addition to (rather than as a mechanisms that earlier evo That is, the different types of ones but are layered on top structures as well, see Badeuo but provocative book Kindi overlapping short-leash an by labelling them as differ the case of humans—and all One key distinction between various systems code for the four different kinds of mind: Popperian mind, and Gregor aly powerful mechanismism (1991) notes, brains are ant reflect increasingly sophistic that the minds, in the orde direct genetic control.

The different minds cont stimul in the environment, thus produces hardwired p “said” metaphorically “do Skinnerian mind uses operable environment (the genes you go along”). The Popper represent possibilities and 1 have “said” metaphorically
and the goal structures keyed into System 1 are given below in Table 7.3.

can only do their best in *advance* by building a fast executive computer for
tics of the living trade . . . The advantage of this sort of programming is that it
greatly cuts down on the number of detailed rules that have to be built into
the original program" (Dawkins, 1976, pp. 55, 57). Human brains represent,
according to Dawkins (1976, pp. 59–60) "the culmination of an evolutionary
trend towards the emancipation of survival machines as executive decision-

makers from their ultimate masters, the genes . . . By dictating the way sur-
vival machines and their nervous systems are built, genes exert ultimate
power over behavior. But the moment-to-moment decisions about what to do
next are taken by the nervous system. Genes are the primary policy-makers;
brains are the executives . . . The logical conclusion to this trend, not yet
reached in any species, would be for the genes to give the survival machine a
single overall policy instruction: do whatever you think best to keep us alive".

The type of long-leash control that Dawkins is referring to is built in
addition to (rather than as a replacement for) the short-leash genetic control
mechanisms that earlier evolutionary adaptation has installed in the brain.
That is, the different types of brain control that evolve do not replace earlier
ones but are layered on top of them (and of course perhaps alter the earlier
structures as well, see Badcock, 2000, pp. 27–29). Dennett (1996), in his short
but provocative book *Kinds of Minds* (see also, Dennett, 1975), describes the
overlapping short-leashed and long-leashed strategies embodied in our brains
by labelling them as different "minds"—all lodged within the same brain in
the case of humans—and all simultaneously operating to solve problems.

One key distinction between Dennett's kinds of minds is how directly the
various systems code for the goals of the genes. Dennett (1996) distinguishes
four different kinds of minds, the Darwinian mind, the Skinnerian mind, the
Popperian mind, and Gregorian mind (Fig. 7.1). The minds reflect increasingly
powerful mechanisms for predicting the future world. As Dennett
(1991) notes, brains are anticipation machines. The four minds he proposes
reflect increasingly sophisticated modes of anticipation. It will be argued here
that the minds, in the order listed above, also reflect decreasing degrees of
direct genetic control.

The different minds control in different ways how the vehicle will react to
stimuli in the environment. The Darwinian mind uses prewired reflexes and
thus produces hardwired phenotypic behavioural patterns (the genes have
"said" metaphorically "do this when x happens because it is best"). The
Skinnerian mind uses operant conditioning to shape itself to an unpredict-
able environment (the genes have "said" metaphorically "learn what is best as
you go along"). The Popperian mind (after the philosopher Karl Popper) can
represent possibilities and test them internally before responding (the genes
have "said" metaphorically "think about what is best before you do it").
Gregorian mind (after the psychologist Richard Gregory) exploits the mental tools (see Clark, 1997) discovered by others (the genes have “said” metaphorically “imitate and use the mental tools used by others to solve problems”). In humans, all four “minds” are simultaneously operative (Fig. 7.1).

The Darwinian and Skinnerian minds have short-leash goals installed (“when this stimulus appears, do this”). In contrast, the Popperian and Gregorian minds are characterized by long-leash goals (“operate with other agents in your environment so as to increase your longevity”).

When confronted with a problem, all these parts of the brain contribute potential solutions. It is variable which one will dominate. We have argued (Stanovich, 1999; Stanovich & West, 2000) that measures of psychometric intelligence are measures of current computational capacity instantiated at the algorithmic level of System 2. This computational capacity is available to be deployed in a System 1 override function if the intentional-level goals of System 2 dictate that this will achieve goal maximization (Fig. 7.2). This override of System-1-triggered responses will not always be successful and thus it is predicted that on tasks where Systems 1 and 2 are triggering different responses, the instrumentally optimal response will be made by individuals with higher psychometric intelligence. It is precisely this that accounts for the pattern of results we reviewed earlier. In short, we argue that high analytic intelligence may lead to task construals that track instrumental rationality; whereas the alternative construals of subjects low in analytical intelligence (and hence more dominated by System 1 processing) might be more likely to track evolutionary rationality in situations that put the two...


WHERE EVOLUTIONARY PSYCHOLOGY GOES WRONG

Consider the bee. As a Darwinian creature, it has a goal structure as indicated in Fig. 7.3. The area labelled A indicates the majority of cases where the

Figure 7.3 Goal structure of a Darwinian creature. The areas indicate overlap and nonoverlap of vehicle and genetic "interests".
replicator and vehicle goals coincide. Not flying into a brick wall serves both
the interests of the replicators (the bee has a function in the hive that will
facilitate replication) and of the bee itself as a coherent organism. Of course,
the exact area represented by A is nothing more than a guess. The important
point is the existence of a nonzero area B—a set of goals that serve only the
interests of the replicators and that are antithetical to the interests of the
vehicle itself.4 A given bee will sacrifice itself as a vehicle if there is greater
benefit to the same genes by helping other individuals (for instance, causing
its own death when it loses its sting while protecting its genetically related
hive-queen). There are no conflicting goals in a Darwinian creature. Its goals
are the genes' goals pure and simple. It is just immaterial as far as evolution-
ary rationality is concerned how much genetic goals overlap with vehicle
goals. Perfect rationality for the bee means local fitness optimization for its
genes, because for the bee the only relevant rationality is evolutionary
rationality.

The error that evolutionary psychologists tend to make is that they stop
right there, with an implicit assumption that evolutionary rationality is all
there is; that there is no instrumental rationality (no maximization issue at the
level of the whole organism—the vehicle). Evolutionary psychologists, in
effect, treat humans as if they were bees. This error comes about for two
reasons. First, despite emphasizing in their writings that the EEA was differ-
ent from the modern environment, evolutionary psychologists have been
reluctant to play out the implications of this fact. Second, because of their
advocacy of a strictly modular view of mind and their tendency to eschew
domain-general mechanisms (Cosmides & Tooby, 1992, 1994b; Tooby &
Cosmides, 1992), evolutionary psychologists de-emphasize the utility of the
flexible goal structures of System 2 and the functions of the serial, system-
atically analytic processes carried out by that system. In short, evolutionary
psychologists take issue with the characterization of the algorithmic level of
System 2 (that it can instantiate domain-general procedures), but in doing so
they miss the important function of the flexible goal structure that rides on
top of the algorithmic level of System 2 (at the intentional level of analysis;
see Stanovich, 1999). They are algorithmic-level mechanisms are
meant by theorists to displace the functionality (and implications) of the
intentional level of System 2.

With the advent of the high-Gregorian type), evolution has a
flexible system that is somewhat
Dawkins: "Do whatever you think is beneficial here. The key point is
long-leash goals, and a Popper
large optimization becomes the first time, we have the possibility
7.4. Here, although we have area A and area B (goals serving the
we have a new area, C (again
chapter are pure conjecture), we serve the vehicle's interests by
Why does area C come to
When they started building I
seeing on to a long-leash strategy
"Things will be changing for

---

4 We will continue the practice here of using the metaphorical language about genes having
"goals" or "interests" in confidence that the reader understands that this is a shorthand only. As
Blackmore (1999, p. 5) notes, "the shorthand 'genes want X' can always be spelled out as 'genes
that do X are more likely to be passed on'," but that, in making complicated arguments, the latter
language becomes cumbersome. Thus, we will follow Dawkins (1976, p. 88) in "allowing ours-
elves the licence of talking about genes as if they had conscious aims, always reassuring our-
ourselves that we could translate our sloppy language back into respectable terms if we wanted to." Dawk-
ins points out that this is "harmless unless it happens to fall into the hands of those ill-
equipped to understand it" (Dawkins, 1976, p. 278) and then proceeds to quote a philosopher
shrewdly and pedantically admonishing biologists that genes can't be selfish any more than atoms
can be jealous. We trust, Dawkins' philosopher aside, that no reader needs this pointed out.

---

Figure 7.4 The logic of the goal

---
7. EVOLUTIONARY VERSUS INSTRUMENTAL GOALS

A brick wall serves both function in the hive that will coherent organism. Of course, a than a guess. The important set of goals that serve only the logical to the interests of the is a vehicle if there is greater virtuals (for instance, causing tecting its genetically related Darwinian creature. Its goals immaterial as far as evolutionary goals overlap with vehicle al fitness optimization for its t rationality is evolutionary nd to make is that they stop evolutionary rationality is all (no maximization issue at the olutionary psychologists, in t error comes about for two ings that the EEA was differ arly psychologists have been act. Second, because of their nd their tendency to eschew oby, 1992, 1994b; Tooby & idemphasize the utility of the cions of the serial, systemsm. In short, evolutionary n of the algorithmic level of p procedures), but in doing so goal structure that rides on intentional level of analysis;

see Stanovich, 1999). They are so focused on denying domain generality in algorithmic-level mechanisms (in part because they mistakenly believe that it is meant by theorists to displace the modular mind, see below) that they miss the functionality (and implications for rationality) of the goal structure at the intentional level of System 2.

With the advent of the higher-level System 2 minds (of the Popperian and Gregorian type), evolution has inserted into the architecture of the brain a flexible system that is somewhat like the ultimate long-lease goal suggested by Dawkins: “Do whatever you think best.” But “best for whom?” is the critical question here. The key point is that for a creature with a flexible intelligence, long-lease goals, and a Popperian/Gregorian mind, we have the possibility of genetic optimization becoming dissociated from the vehicle’s goals. For the first time, we have the possibility of a goal structure like that displayed in Fig. 7.4. Here, although we have area A (where gene and vehicle goals coincide) and area B (goals serving the genes’ interests but not the vehicle’s) as before, we have a new area, C (again, the sizes of these areas in all diagrams in this chapter are pure conjecture). In humans, we have the possibility of goals that serve the vehicle’s interests but not those of the genes.

Why does area C come to exist only in creatures with long-lease goals? When they started building Popperian and Gregorian minds, the genes were giving up on the strategy of coding moment-by-moment responses and moving to a long-lease strategy that at some point was the equivalent of saying “Things will be changing too fast out there, brain, for us to tell you exactly

---

**Figure 7.4** The logic of the goal structure in a human.
what to do—you just go ahead and do what you think is best given the general goals (survival, sexual reproduction) that we (the genes) have inserted'. And there is the rub. In long-leash brains, genetically coded goals can be represented only in the most general sense. There is no goal of "mate with person X at 6:57pm on Friday, June 13" but instead "have sex because it is pleasurable". But once the goal has become this general, a potential gap has been created whereby behaviours that might serve the vehicle's goal might not serve that of the genes. We need not go beyond the obvious example of sex with contraception—an act that serves the vehicle's goal of pleasure without serving the genes' goal of reproduction. What is happening here is that the flexible brain is coordinating multiple long-term goals—including its own survival and pleasure goals—and these multiple long-term goals come to overshadow its reproductive goal. From the standpoint of the genes, the human brain can sometimes be like a Mars Explorer run amok. It is so busy coordinating its secondary goals (master your environment, engage in social relations with other agents, etc.) that it sometimes ignores the primary goal of replicating the genes that the secondary ones were supposed to serve.

Ironically, what from an evolutionary design point of view could be considered design defects actually make possible instrumental rationality—optimizing the utility of the person rather than the fitness of subpersonal units called genes. That is, inefficient design from an evolutionary point of view in effect creates the possibility of a divergence between organism-level goals and gene-level goals, which is an implication of Millikan's (1993, p. 67) point that "there is no reason to suppose that the design of our desire-making systems is itself optimal. Even under optimal conditions these systems work inefficiently, directly aiming, as it were, at recognizable ends that are merely roughly correlated with the biological end that is reproduction. For example, mammals do not, in general, cease efforts at individual survival after their fertile life is over".

Our framework integrates the insight of the possibility of vehicle/replicator goal mismatch in the direction of the vehicle (although the possibility of area B has been acknowledged for some time, the implications of area C have been incompletely worked out) with some assumptions about the intentional-level properties of Systems 1 and 2 drawn largely from Reber (1992, 1993). The integrated framework is displayed in Fig. 7.5 (of course, the exact size of the areas of overlap are mere guesses). Again, an assumption reflected in the figure is that, in the vast majority of real-life situations, evolutionary rationality also serves the goals of instrumental rationality. But the most important feature of the figure is that it illustrates the asymmetries in the "interests" served by the goal distribution of the two systems. The remnants of the Darwinian creature structure (see Fig. 7.3) are present in the System 1 brain structures of humans. Many of the goals instantiated in this system were acquired nonreflectively—they have not undergone an evaluation in terms of whether the fact has been evaluated, but by a sense of the individual person (the very vestiges of our animal past; it is the problem with moral values. We often call an impervious to persuasion, or outcomes are not malfunc-

Figure 7.5 Genetic and vehicle go

7. EVOLUTION
hat you think is best given the
on) that we (the genes) have
h brains, genetically coded goals
sense. There is no goal of "mate
1 but instead "have sex because it
me this general, a potential gap
ght serve the vehicle's goal might
beyond the obvious example of
the vehicle's goal of pleasure.
What is happening here is
long-term goals—including its
multiple long-term goals come to
the standpoint of the genes, the
Explorer run amok. It is so busy
environment, engage in social
imes ignores the primary goal of
were supposed to serve.
ign point of view could be con-
ible instrumental rationality—
than the fitness of subpersonal
(from an evolutionary point of
ence between organism-level
ation of Millikan's (1993, p. 67)
the design of our desire-making
conditions these systems work
cognizable ends that are merely
is reproduction. For example,
t individual survival after their
of the possibility of vehicles/
ve vehicle (although the possibil-
time, the implications of area
some assumptions about the
2 drawn largely from Reber
played in Fig. 7.5 (of course, the
guesses). Again, an assumption
ity of real-life situations, ev-
strumental rationality. But the
it illustrates the asymmetries
ution of the two systems. The
(see Fig. 7.3) are present in
many of the goals instantiated
—they have not undergone an
evaluation in terms of whether they served the person's interests. They have in
fact been evaluated, but by a different set of criteria entirely: whether they
enhanced the longevity and fecundity of the replicators. From the standpoint
of the individual person (the vehicle) these are the dangerous goals, the ones
that sacrifice the vehicle to the goals of replicators—the ones that lead the bee
to sacrifice itself for its genetically related queen. As Pinker (1997, p. 370)
notes, "the problem with emotions is not that they are untamed forces
vestiges of our animal past; it is that they are designed to propagate copies
of the genes that built them rather than to promote happiness, wisdom, or
moral values. We often call an act 'emotional' when it is harmful to the social
group, damaging to the actor's happiness in the long run, uncontrollable and
impervious to persuasion, or a product of self-delusion. Sad to say, these
outcomes are not malfunctions but precisely what we would expect from
well-engineered emotions".

What the right side of Fig. 7.5 (indicating the goal structure of System 2)
depicts is that a bee with a Popperian/Gregorian intelligence might well
decide that it would rather forgo the sacrifice! It is the reflective processes
embodied in System 2 that derive the flexible long-leash goals that often have
utility for the organism but thwart the goals of the genes (sex with contracep-
tion, resource use after the reproductive years have ended, etc.). These are the
goals at the top of the right side of Fig. 7.5 that overlap with vehicle interests
but not genetic interests.

Figure 7.5 Generic and vehicle goal overlap in the two systems.
Failure to acknowledge the divergence of "interests" (see footnote 4, p. 190) between replicators and their vehicles is an oversight that sociobiologists were certainly guilty of (see Symons, 1992, on the "genetic fallacy") and that evolutionary psychologists are sometimes guilty of. For example, evolutionary psychologists are fond of pointing to the optimality of cognitive functioning—of showing that certain reasoning errors that cognitive psychologists have portrayed as a characteristic and problematic aspect of human reasoning (Kahneman & Tversky, 1984, 1996, 2000) have in fact a logical evolutionary explanation (Brase, Cosmides, & Tooby, 1998; Cosmides & Tooby, 1996; Gigerenzer, 1996b; Rode et al., 1999). The connotation, or unspoken assumption, is that therefore there is nothing to worry about—that since human behaviour is optimal from an evolutionary standpoint, the concern for cognitive reform that has been characteristic of many cognitive psychologists (termed "Mellorists" by Stanovich, 1999) has been misplaced. But this sanguine attitude too readily conflates genetic optimization with goal optimization for the vehicle. To avoid the error, the different "interests" of the replicators and vehicles must be recognized—and we must keep evaluations of efficiency consistent with the entity whose optimization is at issue. The bee, as a Darwinian creature, needs no cognitive reform because it has no "interests" other than its genes' interests. Humans, with Gregorian minds, have interests as vehicles and thus might benefit from cognitive reform in situations where vehicle interests conflict with genetic interests and their Darwinian minds are siding with the latter. In such a case, it is imperative that System 2 carry out its override function (as depicted in Fig. 7.2), and suppress the System 1 response and substitute one more congruent with vehicle well-being.

Situations where evolutionary and instrumentality dissociate might well be rare, but the few occasions on which they occur might be important ones. This is because knowledge-based, technological societies often put a premium on abstraction and decontextualization, and they sometimes require that the fundamental computational bias of human cognition toward contextualization of problems (see Stanovich, 1999, in press; Stanovich & West, 2000) be overridden by System 2 processes.

Evolutionary psychologists are prone to emphasize situations where genetic goals and personal goals coincide. They are not wrong to do so, because this is most often the case. Accurately navigating around objects in the natural world was adaptive during the EEA, and it similarly serves our personal goals as we carry out our lives in the modern world. Likewise, with other evolutionary adaptations: It is a marvel that humans are exquisite frequency detectors (Hasher & Zacks, 1979), that they infer intentionality with almost supernatural ease (Levinson, 1995), and that they acquire a complex language code from impoverished input (Pinker, 1994). All of these mechanisms serve personal goal fulfillment in the modern world. But none of this means that the overlap is necessarily 100 per cent.

Unfortunately, the modern world has the default values of evolutionarily normal. Modern technological societies must decontextualize information (Adler, 1984) and in a depersonalized, the context-specific way assumed by the thesis (Samuels, 1998). Such situations personalizing and contextualizing some computational biases (Stanovich, 1990), with the demands for decontextualization by society puts on its citizens. Indeed, the need to teach such skills of cognitive decontextualization is demanding such skills (Dickson & Gottfredson, 1997; Hunt, 1995; McClelland, Wethington, Moen, & Schwartz, 1995). The example, many aspects of the context of detaching prior belief and world views. There has been understated because of the utilitarian theories and not that has nothing to do with the context of background knowledge and personal.

The need to decontextualize also be necessary. Consider the case of "the customer is always right," to include even instances where customer is always right, it must be set aside by the service worker. The customer is always right" must be in the domain of the market-based transmeta she is not not in an actual social interac this call for blocking them on the node realm where different rules apply. Concerns about the real-world necessary cognitive abstraction (to warn against minimizing the real world styles. In discussing the systems of cognitive operations involve some basic not of codes
"interests" (see footnote 4, an oversight that sociobiolo-
 gist, or the "genetic fallacy")
 may be guilty of. For example, the optimality of cognitive
 errors that cognitive pseu-
 dohistoric aspect of human
 2000) have in fact a logical
 Tooby, 1998; Cosmides &
 1999). The connotation, or
 thing to worry about— that
 tory standpoint, the con-
 cerpist of many cognitive
 , 1999) has been misplaced.
 ntic optimization with goal
 be different "interests" of the
 we must keep evaluations of
 zation is at issue. The bee, as
 because it has no "interests"
 gorian minds, have interests
 reform in situations where
 the Darwinian minds are
 System 2 carry out its
 press the System 1 response
 well-being.
 ent rationality dissociate
 which they occur might be
 sed, technological societies
 ontexialization, and they
 tional bias of human cog-
 ; Stanovich, 1999, in press;
 1 processes.
 mphasize situations where
 y are not wrong to do so,
 avaging around objects in
 it similarly serves our
 xandian world. Likewise, with
 it humans are exquisite fre-
 nce infer intentionality with
 hat they acquire a complex
 1994). All of these mechan-
 sm world. But none of this

 Unfortunately, the modern world tends to create situations where some of
 the default values of evolutionarily adapted cognitive systems are not opti-
 mal. Modern technological societies continually spawn situations where
 humans must decontextualize information—where they must deal abstractly
 (Adler, 1984) and in a depersonalized manner with information rather than in
 the context-specific way assumed by proponents of the massive modularity
 thesis (Samuels, 1998). Such situations require the active suppression of the
 personalizing and contextualizing styles that characterize the fundamental
 computational biases (Stanovich, 1999, in press). Such biases directly conflict
 with the demands for decontextualization that a highly bureaucratized soci-
 ety puts on its citizens. Indeed, this is often why schools have to explicitly
 teach such skills of cognitive decontextualization. Increasingly, modern soci-
 ety is demanding such skills (Dickens & Flynn, 2001; Frank & Cook, 1995;
 Gottfredson, 1997; Hunt, 1995, 1999) and, in some cases, it is rendering
 economically superfluous anyone who does not have them (Bronfenbrenner,
 McClelland, Wethington, Moen, & Ceci, 1996; Frank & Cook, 1995). For
 example, many aspects of the contemporary legal system put a premium on
 detaching prior belief and world knowledge from the process of evidence
 evaluation. There has been understandable vexation at odd jury verdicts ren-
 dered because of jury theories and narratives concocted during deliberations
 that had nothing to do with the evidence but that instead were based on
 background knowledge and personal experience.

 The need to decontextualize also characterizes many work settings in con-
 temporary society. Consider the common admonition in the retail service
 sector of "the customer is always right". This admonition is often interpreted
 to include even instances where customers unleash unwarranted and some-
 times astonishingly vitriolic verbal assaults. The service worker is supposed to
 remain polite and helpful under this onslaught, despite the fact that such
 emotional social stimuli are no doubt triggering evolutionarily instantiated
 modules of self defence and emotional reaction. All of this emotion, all of
 these personalized attributions—all fundamental computational biases—
must be set aside by the service worker and instead an abstract rule that "the
 customer is always right" must be invoked in this special, socially constructed
 domain of the market-based transaction. The worker must realize that he or
 she is not in an actual social interaction with this person (which, if true, might
 call for socking them on the nose!), but in a special, indeed "unnatural",
 realm where different rules apply.

 Concerns about the real-world implications of the failure to engage in
 necessary cognitive abstraction (see Adler, 1984) were what led Luria (1976)
 to warn against minimizing the importance of decontextualizing thinking
 styles. In discussing the syllogism he notes that "a considerable proportion
 of our intellectual operations involve such verbal and logical systems; they com-
 prise the basic network of codes along which the connections in discursive
human thought are channeled" (Luria, 1976, p. 101). Einhorn and Hogarth (1981) highlight the importance of decontextualized environments in their discussion of the optimistic and pessimistic views of the cognitive biases revealed in laboratory experimentation. Einhorn and Hogarth (1981, p. 82) note that "the most optimistic asserts that biases are limited to laboratory situations which are unrepresentative of the natural ecology", but they go on to caution that "in a rapidly changing world it is unclear what the relevant natural ecology will be. Thus, although the laboratory may be an unfamiliar environment, lack of ability to perform well in unfamiliar situations takes on added importance" (p. 82).

Critics of the abstract content of most laboratory tasks and standardized tests have been misguided on this very point. Evolutionary psychologists have singularly failed to understand the implications of Einhorn and Hogarth's (1981) warning. They regularly bemoan the "abstract" problems and tasks in the heuristics and biases literature and imply that since these tasks are not like "real life" we need not worry that people do poorly on them. The issue is that, ironically, the argument that the laboratory tasks and tests are not like "real life" is becoming less and less true. "Life", in fact, is becoming more like the tests! Try using an international automated cash dispenser with which you are unfamiliar; or try arguing with your health maintenance organization/insurance company about a disallowed medical procedure. In such circumstances, we invariably find out that our personal experience, our emotional responses, our stimulus-triggered intuitions about social justice are all worthless—all are for naught when talking over the phone to the representative looking at a computer screen displaying a spreadsheet with a hierarchy of branching choices and conditions to be fulfilled. The social context, the idiosyncrasies of individual experience, the personal narrative—all are abstracted away as the representatives of modernist technological-based services attempt to "apply the rules".

Modern mass communication technicians have become quite skilled at implying certain conclusions without actually stating those conclusions (for fear of lawsuits, bad publicity, etc.). Advertisements rely on the fundamental computational bias (particularly its enthymematic processing feature) to fill in the missing information. Margolis (1987; see also Margolis, 1996) warns of the ubiquity of this situation in modern society: "We can encounter cases where the issue is both out-of-scale with everyday life experience and contains important novelties, so that habitual responses can be highly inappropriate responses. The opportunity for unrecognized contextual effects akin to the scenario effects ... [demonstrated in the laboratory] can be something much more than an odd quirk that shows up in some contrived situation" (Margolis, 1987, p. 168).

Evolutionary psychologists have argued that some problems can be solved more efficiently if represented to coincide with how various brain modules represent information. Never the world will not always let us suited to our evolutionary de- series of elegant experiments, have shown how at least part bration studies is due to experiments—stimuli that do not, which are optimally suited instances in real life when we cue validity has changed. A strategies for suppressing in automatic responses to cues v who aspires to a career in my university and sees large nume they persist in their old con different specialty area when ations where accomplishment gent performance requirements the competitive environment we are in a situation just l experiments with their unreq learned strategies that will t (Koriat, Lichtenstein, & Fisl HOW EVOL

Dawkins (1976, p. 234) note gene and individual body as psychologists have missed t between the evolutionary of humans are vehicles with int Humans aspire to be more th their genes (which are replica

---

1 Frequency representations of it is claimed that frequentist repre this claim remains controversial. O that the favourable evidence with n the use of problems with transparencies sometimes attenuate cognitive Perham, Over, & Thompson, 2000; 1998; Mellor et al., 2001; Stoman, also Chapter 6).
represent information. Nevertheless, they often seem to ignore the fact that the world will not always let us deal with representations that are optimally suited to our evolutionarily designed cognitive mechanisms. For example, in a series of elegant experiments, Gigerenzer, Hoffrage, and Kleinbolting (1991) have shown how at least part of the overconfidence effect in knowledge calibration studies is due to the unrepresentative stimuli used in such experiments—stimuli that do not match the participants' stored cue validities, which are optimally tuned to the environment. But there are many instances in real life when we are suddenly placed in environments where the cue validities have changed. Metacognitive awareness of such situations and strategies for suppressing incorrect confidence judgements generated by automatic responses to cues will be crucial here. Every high school musician who aspires to a career in music has to recalibrate when he or she arrives at university and sees large numbers of talented musicians for the first time. If they persist in their old confidence judgements they may not change to a different specialty area when this would be advisable. Many real-life situations where accomplishment yields a new environment with even more stringent performance requirements share this logic. Each time we “ratchet up” in the competitive environment of a capitalist economy (Frank & Cook, 1995) we are in a situation just like the overconfidence knowledge calibration experiments with their unrepresentative materials. It is important to have learned strategies that will temper one’s overconfidence in such situations (Korsia, Lichtenstein, & Fischhoff, 1980).

**HOW EVOLUTIONARY PSYCHOLOGY GOES WRONG**

Dawkins (1976, p. 234) notes that there is an “uneasy tension ... between gene and individual body as fundamental agent of life”. Many evolutionary psychologists have missed this essential tension by focusing on parallels between the evolutionary optimization of humans and other animals. But humans are vehicles with interests beyond those of their genes’ replication. Humans aspire to be more than mere survival machines serving the “ends” of their genes (which are replication pure and simple). Only humans really turn

---

3 Frequency representations of probabilistic information are one example. However, although it is claimed that frequentist representations can eliminate cognitive illusions (Gigerenzer, 1991), this claim remains controversial. Over (Chapter 6), for example, presents evidence suggesting that the favourable evidence with respect to frequentist representations may have resulted from the use of problems with transparent logical forms. Furthermore, even if frequency representations sometimes attenuate cognitive illusions, they do not remove them entirely (Evans, Handley Perham, Over, & Thompson, 2000; Giroto & Gonzalez, 2001; Harris & Harvey, 2000; Macchi, 1998; Mellers et al., 2001; Sloane, Over, & Slovak (in press); Tversky & Kahneman, 1983; see also Chapter 6).
the tables (or at least have the potential to) by occasionally ignoring the
terests of the genes in order to further the interests of the vehicle. Evo-
lutionary psychology, for all its important insights into the functioning
human beings, failed to develop this profound insight. By failing to highlight the
interests of the vehicle in discussions of optimal cognitive functioning, evo-
lutionary psychology has colluded with the genes in delivering their most
sophisticated vehicle (human beings) over to them, as if this vehicle—like the
bee—had no interests other than replication.

As argued above, evolutionary psychologists relegate the evolutionary/in-
strumental rationality distinction to the background because many are: (1)
wedded to a cognitive architecture that displays massive modularity; (2) as a
result, they eschew domain-general System 2 mechanisms; and (3) they con-
join these two theoretical assumptions with a tendency to ignore the implica-
tions of mismatches between the EEA and the cognitive requirements of
technological societies.

To the extent that modern society increasingly requires the fundamental
computational biases to be overridden, then dissociations between evolution-
ary and individual rationality will become more common, and System 2 over-
rides will be more essential to personal well-being. Cosmides and Tooby
(1996, p. 15) argue that "in the modern world, we are awash in numerically
expressed statistical information. But our hominid ancestors did not have
access to the modern accumulation which has produced, for the first time in
human history, reliable, numerically expressed statistical information about
the world beyond individual experience. Reliable numerical statements about
single event probabilities were rare or nonexistent in the Pleistocene". "It is
easy to forget that our hominid ancestors did not have access to the modern
system of socially organized data collection, error checking, and information
accumulation... In ancestral environments, the only external database avail-
able from which to reason inductively was one's own observations" (Brase et
al., 1998, p. 5).

Although this may be entirely correct (but see footnote 5, p. 197), let us
carry through with the implications of this point. We are living in a techno-
lological society where we must: decide which health maintenance organization
to join based on just such statistics; figure-out whether to invest in an indi-
vidual retirement account or personal pension plan; decide what type of
mortgage to purchase; figure-out what type of insurance policy to buy; decide
whether to trade-in a car or sell it ourselves; decide whether to lease or to buy;
think about how to apportion our retirement funds; and decide whether we
would save money by joining a book club—to simply list a random set of the
plethora of modern-day decisions and choices. And we must make all of
these decisions based on information represented in a manner for which our
brains may not be adapted (in none of these cases have we coded individual
frequency information from our own personal experience). To reason

normatively in all of these does going to have to deal with p
quentistic terms—in represent argued are different from our
information (Cosmides & Tooby).

Consider the work of Bras dif
probability problem ( 1995 by presenting the info
objects—both alterations d
computation systems of the
adequate performance obs
contain such well-designed
(1998, p. 13) replied that "in
paper-and-pencil problems i
operate on actual events
postindustrial society, we an
the time, and much of what
the perception of actual ever
prepackaged, and condense
percentages, tables, and gr
routinely presented in USA)

What are we attempting t
discussions of findings in ev
ognition literature as well (s
nothing to be gained from
abstract level (the conjunctic
flexibly overriding the fund
 tendency of evolutionary psy
statement (Tooby & Cosmid

In actuality, adaptationist ap
unity of humankind is genu
applies in a privileged way
complexly organized dimen
ences among humans that at
found are so overwhelming
functionally superficial prop

This statement provides an e
goes off the rails. Let us see
ecologists have found" and
reflect "functionally superfic
Well, for starters, some
t) by occasionally ignoring the interests of the vehicle. Evosights about human behaviour, t. By failing to highlight the nal cognitive functioning, evo- genes in delivering their most hem, as if this vehicle—like the 
ists relegate the evolutionary/ kground because many are: (1) as massive modularity; (2) as a mechanisms; and (3) they con- endency to ignore the implica- he cognitive requirements of 
gly requires the fundamental sociations between evolution- common, and System 2 over- being. Cosmides and Tooby . we are awash in numerically nimid ancestors did not have produced, for the first time in statistical information about e numerical statements about ent in the Pleistocene". "It is ot have access to the modern or checking, and information only external database avail- own observations" (Brase et 
nee footnote 5, p. 197), let us it. We are living in a techno- th maintenance organization whether to invest in an indi- plan; decide what type of insurance policy to buy; decide le whether to lease or to buy; nds; and decide whether we rely list a random set of the And we must make all of l in a manner for which our es have we coded individual al experience). To reason 
normatively in all of these domains (to maximize our personal utility) we are going to have to deal with probabilistic information represented in nonfrequentistic terms—in representations that the evolutionary psychologists have argued are different from our adapted algorithms for dealing with frequency information (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995).

Consider the work of Brase et al. (1998), who improved performance on a difficult probability problem (Bar-Hillel & Falk, 1982; Falk, 1992; Granberg, 1995) by presenting the information as frequencies and in terms of whole objects—both alterations designed to better fit the posited frequency-computation systems of the brain. In response to a query about why the adequate performance observed was not even higher given that our brains contain such well-designed frequency-computation systems, Brase et al. (1998, p. 13) replied that "in our view it is remarkable that they work on paper-and-pencil problems at all. A natural sampling system is designed to operate on actual events". The problem is that in a symbol-oriented postindustrial society, we are presented with paper-and-pencil problems all the time, and much of what we know about the world comes not from the perception of actual events but from abstract information preprocessed, prepackaged, and condensed into symbolic codes such as probabilities, percentages, tables, and graphs (the voluminous statistical information routinely presented in USA Today and Social Trends comes to mind).

What we are attempting to combat here is a connotation implicit in some discussions of findings in evolutionary psychology and indeed in the situated cognition literature as well (see Anderson, Reder, & Simon, 1996) that there is nothing to be gained from being able to understand a formal rule at an abstract level (the conjunction rule of probability, etc.), and no advantage in flexibly overriding the fundamental computational biases. We can see the tendency of evolutionary psychologists to fall into this trap in the following statement (Tooby & Cosmides, 1992, p. 79):

In actuality, adaptationist approaches offer the explanation for why the psychic unity of humankind is genuine and not just an ideological fiction; for why it applies in a privileged way to the most significant, global, functional, and complexly organized dimensions of our architecture; and for why the differences among humans that are caused by genetic variability that geneticists have found are so overwhelmingly peripheralized into architecturally minor and functionally superficial properties.

This statement provides an example of how and why evolutionary psychology goes off the rails. Let us see what is in some of that "genetic variability that geneticists have found" and let us ask ourselves, seriously, whether it does reflect "functionally superficial properties".

Well, for starters, some of that "genetic variability that geneticists have
found" is in general intelligence (g), which virtually everyone who has looked at the evidence agrees is at least 40–50 per cent heritable (Deary, 2000; Grigorenko, 1999; Neisser et al., 1996; Plomin & Petrill, 1997). Is g a "functionally superficial" individual difference property of human cognition? No responsible psychologist thinks so. It is, indeed, the single most potent psychological predictor of human behaviour in both laboratory and real-life contexts that has ever been identified (Lubinski, 2000; Lubinski & Humphreys, 1997). It is a predictor of real-world outcomes that are critically important to the maximization of personal utility (to instrumental rationality) in a modern technological society. Objective measures of the requirements for cognitive abstraction have been increasing across most job categories in technological societies throughout the past several decades (Gottfredson, 1997). This is why measures of the ability to deal with abstraction such as g remain the best employment predictor and the best earnings predictor in postindustrial societies (Brody, 1997; Gottfredson, 1997; Hunt, 1995). The psychometric literature contains numerous indications that cognitive ability is correlated with the avoidance of harmful behaviours and with success in employment settings, as well as social status attainment (MacDonald & Geary, 2000), independent of level of education (Brody, 1997; Gottfredson, 1997; Hunt, 1995; Lubinski & Humphreys, 1997).

We view individual differences in g as indicating differences in the current computational capacity of the algorithmic level of System 2. It is critically related to the override function of System 2 discussed above—the override function necessary to trump the fundamental computational biases of System 1 when they lead to a response that is antithetical to the interests of the vehicle.

Our algorithmic-level understanding of System 2 borrows from Dennett (1991), who conceives of System 2 as a serial von Neumann computer simulated by the massively parallel computational network of the brain. It is language-based, rule-based, and at least more logic-based than System 1 (Evans & Over, 1996, 1997)—and is the focus of our awareness (it is the system we use to construct a model of the self). As mentioned previously, Evans and Over (1999) discuss the fundamental importance of System 2 as the mechanism that supports hypothetical thinking. In contrast to the holistic/associative nature of System 1, System 2 is analytic in operation, and it is demanding in terms of computational capacity. We view general intelligence to encompass two fundamental classes of property (that perhaps map into the fluid/crystallized distinction from the Horn/Cattell model, Horn, 1982; Horn & Cattell, 1967). First, there is the computational power of the parallel network to sustain the serial simulation* (this is probably closer to fluid intelligence in the Horn/Cattell model). The major factor is the power of the Gregorian mind in Dennett's words. Differences in this factor might be a reflection of the Horn/Cattell model).

Intelligence is not the only concept that could be told about many people; it could be told about many people at all levels of cognitive variability, see the raiser heritable but also important factors noted above. For example, rationality psychologists against the view of Tooby and Cosmides' differences are to species-typical the colors of the wires in a components", and points to such such as work, status attain prophetic relationships to be for the individual (job success) for substance abuse.

Despite Buss's (1999) and other influential evolutionary the psychological processes (and presumably lack any important to the vehicle):

* Deary (2000) has written a book summarizing the evidence on the relationship between reaction time and other speeded tasks and intelligence. It turns out to be difficult to explain why
virtually everyone who has 0 per cent heritable (Deary, Loewy & Petrill, 1997). It is a property of human cognition, indeed, the single most valued in both laboratory and ecological tasks (Lubinski, 2000; Lubinski & Benigeri, 1994) that are critical to individual utility (to instrumental goals). Objective measures of the variables, increasing across most jobs in the past several decades (some ability to deal with abstractions and the best earnings) are sensitive to the current level of education (Brody & Humphreys, 1997). Differences in the current 1 of System 2. It is critically discussed above—the override computational biases of System 1 to the interests of the stem 2 borrows from Dennett’s von Neumann computational network of the brain. It is logic-based than System 1 of our awareness (it is the I). As mentioned previously, 1 importance of System 2 as linking. In contrast to the innate, analytic in operation, and occur. We view general intelligence (property that perhaps map Horn/Cattell model, Horn, computational power of the n5 (this is probably closer to fluid intelligence in the Horn/Cattell model of intelligence). The second major factor is the power of the cultural tools used during serial simulation—the Gregorian mind in Dennett’s (1991) Tower of Intellect model (individual differences in this factor might relate to variance in crystallized intelligence in the Horn/Cattell model).

Intelligence is not the only type of "genetic variability that geneticists have found" that is manifestly not "functionally superficial". Similar stories could be told about many personality variables (reflective of intentional level cognitive variability, see Stanovich, 1999) that have been shown to be heritable but also important predictors of behaviour (see p. 394 of Buss, 1999; Matthews & Deary, 1998). Indeed, this stance by some evolutionary psychologists against heritable cognitive traits with demonstrable linkages to important real-world behaviours has become an embarrassment even to some evolutionary theorists. Buss (1999, p. 394) characterizes the view of Tooby and Cosmides as the notion that "heritable individual differences are to species-typical adaptations, in this view, as differences in the colors of the wires in a car engine to the engine's functional working components", and points to some of the same embarrassing empirical facts noted above. For example, heritable personality traits such as conscientiousness and impulsivity have been related to important life goals such as work, status attainment, mortality, and faithfulness in partnerships. Buss’s (1999) alternative interpretation is in terms of genetic concepts such as frequency-dependent selection. But whether or not one accepts such explanations, the point is that many evolutionary theorists have mistakenly downplayed cognitive constructs that are heritable (intelligence, personality dimensions, thinking styles) and that have demonstrated empirical relationships to behaviours that relate to utility maximization for the individual (job success, personal injury, success in relationships, substance abuse).

Despite Buss’s (1999) more nuanced position on individual differences, other influential evolutionary psychologists repeat like a mantra the view that any psychological processes with genetic variation lack any importance (and presumably lack any relevance for rationality, since this is obviously important to the vehicle):

elementary information processing tasks correlate with intelligence at all. Stanovich (2001) conjectured that it is not because they measure some inherent "mental speed" (Deary reviews evidence indicating that the RT-IQ relationship is virtually unchanged when differences in nerve conduction speed are partialled out). But the speed component of these IQ tasks may not be the critical thing. Rather, they all may serve as indirect indicators of the computational power available in the brain’s connectionist network—computational power that is available to sustain the simulation of a serial processor. Of course, there are other more direct indicators of the computational power available to sustain serial simulation, such as working memory, and not surprisingly these indicators show larger correlations with intelligence.
Human genetic variation ... is overwhelmingly sequestered into functionally superficial biochemical differences, leaving our complex functional design universal and species typical. (Tooby & Cosmides, 1992, p. 25)

Humans share a complex, species typical and species-specific architecture of adaptations, however much variation there might be in minor, superficial, nonfunctional traits. (Tooby & Cosmides, 1992, p. 38)

One boggles at general intelligence—one of the most potent psychological predictors of life outcomes—being termed "nonfunctional". But then one realizes what is motivating these statements—a focus on the gene. Even if one buys the massive-modularity-of-adaptations line of the evolutionary psychologist and views general intelligence as some kind of spandrel or byproduct, 7 from the standpoint of the vehicle's interests, it is certainly not nonfunctional. Only a focus on the subpersonal replicators would spawn such a statement—one that relegates important cognitive traits such as intelligence and conscientiousness to the background (Lubinski, 2000; Matthews & Deary, 1998). As soon as one focuses on the organismic level of optimization rather than genetic optimization, the "nonfunctional" traits spring to the foreground as the System 2 algorithmic-level (intelligence) and intentional-level (conscientiousness, openness) constructs that explain individual differences in attaining one's goals (Baron, 1993, 1994; Stanovich, 1999).

The downplaying of the importance of a heritable cognitive indicator such as general intelligence by evolutionary psychologists often results from their tendency to caricature cognitive theories that stress a domain-general mechanism (like the type of analytical processing hypothesized for System 2 by many dual-process theorists). The evolutionary theorists purport to dispute theories that view the evolutionary history of human cognition as the replacement of context-dependent modules with context-independent general intelligence mechanisms. For example, in attacking the so-called Standard Social Science Model (SSSM), Tooby and Cosmides (1992, p. 113) argue that this default social science model "views an absence of content-specific structure as a precondition for richly flexible behavior". Their view of the standard cognitive model in psychology is that general processing mechanisms replace domain-specific ones. Actually, as the long history of dual-process models attests (see Table 7.3), the standard view in psychology is

7 Although the outcome of disputes about whether general intelligence is a byproduct or adaptation does not alter our argument, it should be noted that theorists such as LaCerra and Bingham (1998) and Foley (1996) argue that the changing online requirements of the ancestral hominid environment would, unlike the massive modularity thesis, have required a flexible general intelligence (see also, Nocić, 1993, p. 120, for a philosophically oriented version of a similar argument).
much more similar to the evolutionary psychology view than Tooby and Cosmides want to admit. As in Dennett’s (1996) “Tower of Intellect” model, all of the two-process views listed in Table 7.3 conceive of analytical processes developing in conjunction with domain-specific mechanisms (see Mithen, 1996). Analytical processing mechanisms develop in addition to System 1 modules—they do not replace them.

Evolutionary psychologists also tend to misleadingly minimize the consequences of mismatches between the EEA and the modern environment. Tooby and Cosmides (1992, p. 72) approvingly paraphrase Shepard’s (1987) point that evolution insures a mesh between the principles of the mind and the regularities of the world. But this “mesh” concerns regularities in the EEA, not in the modern world—with its unnatural requirements for decontextualization (requirements that do not “mesh” with the fundamental computational biases toward comprehensive contextualization of situations).

One page later in their chapter, Tooby and Cosmides (1992, p. 73) reveal the characteristic bias of evolutionary psychologists—the belief that “often, but not always, the ancestral world will be similar to the modern world (e.g. the properties of light and the laws of optics have not changed)”. We largely agree. However, although the laws of optics haven’t changed, the type of one-shot, abstract, probabilistic, and symbolically represented decision situations a modern human being must deal with are certainly unprecedented in human history. Think of insurance decisions, retirement decisions, investment decisions, home-buying decisions, relocation decisions, and school choices for children. These are not the highly practised, frequency-coded, time-pressured, recognition-based situations (Klein, 1998) where evolutionary heuristics work best. Instead, these are all the types of situation that invoke just the type of representativeness, availability, sunk cost, confirmation bias, overconfidence, and other effects that the heuristics and biases researchers have studied (see the many real-life examples in Kahneman & Tversky, 2000).

We can walk and navigate among objects as well as we ever did, but no evolutionary mechanism has sculpted my brain to estimate the type of insurance policy we need or how we should evaluate the cost of a disability policy to cover salary loss.

Tooby and Cosmides (1992) seem to take a completely one-sided message from the potential mismatch between the EEA and modern conditions, when in fact the mismatch has more than one implication. Using the example of how our colour constancy mechanisms fail under modern sodium vapour lamps, they warn that “attempting to understand color constancy mechanisms under such unnatural illumination would have been a major impediment to progress” (Tooby & Cosmides, 1992, p. 73)—a fair enough point. But our purpose here is to stress a different corollary point that one might have drawn. The point is that if the modern world were structured such that making colour judgements under sodium lights was critical to our well-being,
then this would be troublesome for us because our evolutionary mechanisms have not naturally equipped us for this. One might be given impetus to search for a cultural invention that would circumvent this defect (relative to the modern world, not the EEA) in our cognitive apparatus.

We argue that humans in the modern world are in just this situation vis-à-vis the mechanisms needed for fully rational action in industrial and bureaucratized societies. The processing of probabilistic information provides a case in point. We argued above that it is critical to many tasks faced by a full participant in a First World society. Of course, the heuristics and biases literature is full of demonstrations of the problems that people have in dealing with probabilistic information. Evolutionary psychologists have done important work that suggests that the human cognitive apparatus may be more adapted to dealing with frequencies than with probabilities (Brase et al., 1998; Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995; but see footnote 5 and Chapter 5). For example, it has been found that when tasks such as the Linda Problem, knowledge calibration tasks, and base-rate tasks are revised in terms of estimating the frequency of categories rather than judging probabilities that performance is improved (see Cosmides & Tooby, 1996; Fiedler, 1988; Gigerenzer, 1991, 1993; Gigerenzer & Hoffrage, 1995; Tversky & Kahneman, 1983; but see Mellers, Hertwig, & Kahneman, 2001). As useful as this research has been (and, indeed, it can usefully be adapted to tell us how to more understandably present probabilistic information in real-life settings, see Gigerenzer, Hoffrage, & Ebert, 1998), it will not remove the necessity of being able to process probabilistic information when it is presented in the real world.

The evolutionary psychologists and ecological rationality theorists are sometimes guilty of implying just this—that if the human cognitive apparatus can be shown to have been adapted during evolution to some other representation (other than that required for a problem in modern society) then somehow it has been shown that there really is no cognitive problem. For example, in the titles and subheadings of several papers on frequency representation, Gigerenzer (1991, 1993, Gigerenzer et al., 1991) has used the phrasing “how to make cognitive illusions disappear”. This is a strange way to phrase things, because the original illusion has of course not “disappeared”. As Kahneman and Tversky (1996) note, the Muller–Lyer illusion is removed when the two figures are embedded in a rectangular frame, but this does not mean that the original illusion has “disappeared” in this demonstration (see also Samuels, Stich, & Tremoulet, 1999). The cognitive illusions in their original form still remain (although their explanation has perhaps been clarified by the different performance obtained in the frequency version), and the situations (real-life or otherwise) in which these illusions occur have not been eliminated. Banks, insurance companies, medical, and many other institutions of modern society are still exchanging information using linguistic terms like “probabilistically”.

Our physician may on occasion instance with the assurance that in such a case. As many Bayesian investors today are likely to be the cent of his patients so advised.

Drawing on Sperber’s (1996) the proper domain (the moder (1999, p. 114) argue that “we s who describe Darwinian models that normative evaluation should be those who offer a socially assaying their evaluations to the contains a vast array of info different from anything our P Davies, 1996; Looren de Jong are guilty of some disproportion by emphasizing the proper do the actual domain, and the M prone to emphasize the errors acknowledge that humans rea.

Buss (1999, p. 378) shows the “[i. humans are so riddled with errors and biases, how can it pass any system that can be doing an unpublished paper by made that our criteria for rec parochial]” (Buss, 1999, p. 37 eging of the present environment unnecessarily parochial. Th environment in which we mu times as Buss (1999) repositing out that they occur in “a seems damming to his own a novel symbolic situations are and citizens in technological.

With respect to the “arti trots out the old sodium vap have used “artificial, evolu analogous to sodium vapour (1999) takes exactly the wrong EEA and modern conditions in situations where we must s
7. EVOLUTIONARY VERSUS INSTRUMENTAL GOALS

Our evolutionary mechanisms might be given impetus to search for this defect (relative to the paratus).

Here in just this situation vis-à-vis the heuristic and biases problems that people have in dealing with probabilities (Brase et al., 1995; Hoffrage, 1995; but see footnote 7 that when tasks such as tasks, and base-rate tasks are under categories rather than judging the Cosmides & Tooby, 1996; FMR & Hoffrage, 1995; Tversky & Kahneman, 2001). As useful and usefulfully be adapted to tell us heuristic information in real-life (1998), it will not remove the heuristic information when it is

The rationality theorists are the human cognitive apparatus powerful enough to evolve to some other problem in modern society) there is no cognitive problem. For all papers on frequency representation (Buss et al., 1991) has used the illusion or "disappear". This is a strange way it has of course not "disappear", the Muller-Lyer illusion is visible. The illusion is used in a rectangular frame, but this illusion is "appeared" in this demonstration (1999). The cognitive illusions in the explanation has perhaps been known in the frequency version), and in these illusions occur have analogies, medical, and many changing information using

Linguistic terms like "probability" and applying that term to singular events. Our physician may on occasion give us a migraine prescription (Imitrex, for instance) with the assurance that he is 90 per cent certain it will work in our case. As many Bayesian investigators in the calibration literature have pointed out, it is likely that we would be quite upset if we found out that for 50 per cent of his patients so advised the medication did not work.

Drawing on Sperber’s (1994) distinction between the actual domain and the proper domain (the modern environment versus the EEA), Samuels et al., (1999, p. 114) argue that "we suspect that those Panglossian-inclined theorists who describe Darwinian modules as ‘elegant machines’ are tacitly assuming that normative evaluation should be relativized to the proper domain, while those who offer a blazer maker of human rationality are tacitly relativizing their evaluations to the actual domain, which, in the modern world, contains a vast array of information-processing challenges that are quite different from anything our Pleistocene ancestors had to confront." (see also Davies, 1996; Looren de Jong & van der Steen, 1998). Perhaps both groups are guilty of some disproportionate emphasis here. Evolutionary theorists err by emphasizing the proper domain so much that they seem to forget about the actual domain, and the Meliorists in the heuristics and biases camp are so prone to emphasize the errors occurring in the actual domain that they fail to acknowledge that humans really are optimally designed for a proper domain.

Buss (1999, p. 378) shows the former tendency when he asks the question: "If humans are so siddled with cognitive mechanisms that commonly cause errors and biases, how can they routinely solve complex problems that surpass any system that can be developed artificially?"—and answers it by quoting an unpublished paper by Tooby and Cosmides where the argument is made that our criteria for recognizing sophisticated performance "have been parochial" (Buss, 1999, p. 378). Buss seems to be calling our natural privileging of the present environment—the one we actually have to operate in—unnecessarily parochial. The devaluing of the actual decontextualized environment in which we must operate in modern technological society continues as Buss (1999) repeatedly minimizes rational thinking errors by pointing out that they occur in "artificial or novel" (p. 378) situations. The latter seems damning to his own argument (that these errors are trivial) because novel symbolic situations are exactly what bureaucratically immersed workers and citizens in technological societies must constantly deal with.

With respect to the "artificial situations" criticism, Buss (1999, p. 379) trots out the old sodium vapour lamps example, saying that the experiments have used "artificial, evolutionarily unprecedented experimental stimuli analogous to sodium vapour lamps". Like Tooby and Cosmides (1992), Buss (1999) takes exactly the wrong message from the potential mismatch between EEA and modern conditions. It is a very serious worry that we are essentially in situations where we must work under sodium vapour lamps! The cognitive
equivalent of the sodium vapour lamps are: the probabilities we must deal with; the causation we must infer from knowledge of what might have happened; the vivid advertising examples we must ignore; the unrepresentative sample we must disregard; the favoured hypothesis we must not privilege; the rule we must follow that dictates we ignore a personal relationship; the narrative we must set aside because it does not square with the facts; the pattern that we must infer, which is not there because we know a randomizing device is involved; the sunk cost that must not affect our judgement; the judge's instructions we must follow despite their conflict with common sense; the contract we must honour despite its negative affects on a relative; the professional decision we must make because we know it is beneficial in the aggregate even if unclear in this case. These are all the "sodium vapour lamps" that modern society presents to our cognitive apparatus and, if evolution has not prepared us to deal with them, so much the worse for our rational behaviour in the modern world (Stanovich, 1999, in press). Luckily, the Gregorian tools of rational thought, running as virtual machines on our System 2 serial simulator are there to help us in situations such as this.

THE SLIPPERY NOTION OF ECOLOGICAL RATIONALITY

Many of the foregoing arguments about matching conceptions of rationality to the level of the entity being optimized apply to the concept of ecological rationality as well, and the work of those who have championed this concept. But the concept itself is not straightforward. A textual analysis of its usage reveals that it is a slippery concept indeed.

Typical of these confusions is a statement at the end of a volume summarizing the work of one of the laboratories responsible for popularizing the term: "Ultimately, ecological rationality depends on decision making that furthers an organism's adaptive goals in the physical or social environment" (Gigerenzer & Todd, 1999, p. 364). In statements such as this we see a double ambiguity, which makes the ecological rationality term devilishly difficult to pin down and hence to evaluate. First, the phrasing "organism's adaptive goals" makes it unclear what level of analysis we are talking about. The word "adaptive" suggests we are talking, in the technical sense, about evolutionary (hence genetic) goals—that ecological rationality is about how organisms are optimized to achieve the goals of their genes. On the other hand, one looks at the same phrase and wonders whether the word "organism" is not key here—that the word adaptive is actually being used more colloquially—and that we are to put a stress on it (as in the "organism's adaptive goals") and view the ecological rationality concept as akin to instrumental rationality (as maximizing the vehicle's utility). This ambiguity in the "organism's adaptive goals" phrase introduces a second ambiguity into the second part of the quote. Because we are unsure or the vehicle's goals it becomes an environment" is meant to refer to EEA.

These two interpretations of and out throughout the 400 paq and the Adaptive Behavior and in Berlin (see also, Todd & Gig in 1999) draw tight links between work: iity ("ecological rationality is μ suggest that genetic fitness max-

ity is all about. This view is reinforced by the collection of sp has built into the human mind ing" (p. 30) include the fast and chapters in the Gigerenzer ar that "evolution would seize u such as this one and exploit the decision-making organism an a

Buss (1999), discussing To d rationality in an unpublished study the evolutionary sense: "Over had certain statistical regularities ecological structure. Ecologic contai- ing design features the adaptive problem solving" (B and Todd (1999), seems to c designed to optimize is genetic.

After all this emphasis on ecological rationality sits, it is (2000) reply to a critique of O psychology is grounded in en. Unlike Buss's (1999) emphasis time," Todd et al. (2000, p. 37 quotes above, instead assert th making in present environm

consequences".

In short, there is consid logical rationality theorists ab tion to the EEA or to the π ecological rationality is maxim needs to be explicitly identified rationality (Stanovich, 1999)
the probabilities we must deal with are the kind of what might have happened; the unrepresentative subset of the categories we must not privilege; the personal relationship; the common sense of the facts; or the conflict with common sense; or the nature of the facts; or we know it is beneficial in those cases. The term "sodium vapour" was probably first used in a scientific context, not to denote a particular apparatus. Although the term was introduced into the cognitive literature by Stanovich (1999, in press), it is not widely used in scientific contexts. Fortunately, there are several virtual machines available that can simulate different situations.

7. EVOLUTIONARY VERSUS INSTRUMENTAL GOALS

In conclusion, the idea of rationality is central to the concept of ecological rationality, which has been championed by many philosophers and psychologists. A textual analysis of this concept is presented below.

The end of a volume summarizing evolutionary theory on decision making highlights the importance of the physical or social environment in shaping decision-making processes. As such, it is important to distinguish between different types of rationality, where evolutionary rationality is about how organisms are adapted to their environments. It is clear that the term "organism" is not key here—more colloquially—and that we can talk about adaptive goals. To emphasize the point, the "organism's adaptive goals" are not just any part of the organism, but the part that is most important in the evolutionary sense.

Because we are unsure whether adaptive goals refer to the genes' goals or the vehicle's goals, it becomes unclear whether "the physical or social environment" is meant to refer to the current (modern) environment or to the EEA.

These two interpretations of the phrase "organism's adaptive goals" slip in and out throughout the 400 pages of the book by Gigerenzer and Todd (1999) and the Adaptive Behavior and Cognition Group at the Max Planck Institute in Berlin (see also, Todd & Gigerenzer, 2000). On page 335 of the book, they draw tight links between work in behavioural ecology and ecological rationality ("ecological rationality is what behavioral ecology is all about"). This suggests that ecological fitness maximization is, likewise, what ecological rationality is all about. This view is reinforced in other parts of the book where we are told that "the collection of specialized cognitive mechanisms that evolution has built into the human mind for specific domains of inference and reasoning" (p. 30) include the fast and frugal heuristics that are the focus of a dozen chapters in the Gigerenzer and Todd (1999) book. Elsewhere, we are told that "evolution would seize upon informative environmental dependencies such as this one and exploit them with specific heuristics if they would give a decision-making organism an adaptive edge" (Gigerenzer & Todd, 1999, p. 19).

Buss (1999), discussing Tooby and Cosmides' use of the term ecological rationality in an unpublished manuscript, likewise links the term to fitness in the evolutionary sense: "Over evolutionary time, the human environment has had certain statistical regularities... These statistical regularities are called ecological structure... Ecological rationality consists of evolved mechanisms containing design features that utilize this ecological structure to facilitate adaptive problem solving" (Buss, 1999, p. 378). Thus, Buss, like Gigerenzer and Todd (1999), seeks to clearly imply that what ecological rationality is designed to optimize is genetic fitness in the EEA.

After all this emphasis on evolution being the superstructure on which ecological rationality sits, it is a surprise to hear Todd, Fiddick, and Krauss (2000) reply to a critique of Over (2000) by saying that although evolutionary psychology is grounded in ecological rationality the converse is not true. Unlike Buss's (1999) emphasis on the human environment "over evolutionary time," Todd et al. (2000, p. 379), in a complete theoretical reversal from the quotes above, instead assert that ecological rationality "encompasses decision making in present environments without privileging problems with fitness consequences".

In short, there is considerable inconsistency in the writings of the ecological rationality theorists about whether ecological rationality is optimization to the EEA or to the modern environment—in short, about whether ecological rationality is maximizing for the genes or for the vehicle. The term needs to be explicitly identified with what Stanovich has termed evolutionary rationality (Stanovich, 1999) or with what has traditionally been viewed as...
instrumental rationality—rationality for the whole organism (see Over, 2000). Instead, these theorists seem to slip back and forth in their usage, adopting whichever stance is most convenient for the argument being made. Ecological rationality theorists seem to want the imprimatur of evolution (and the biological plausibility that evolutionary adaptation provides) without accepting other inconvenient implications of evolutionary explanations. One implication is that we cannot assume that System 1 heuristics (adapted for the EEA) are optimal for achieving rationality in the modern world (Stanovich, 1999, in press). Many important decisions in life are nearly “one shot” affairs (job offers, pension decisions, investing decisions, housing decisions, marriage decisions, reproductive decisions, etc.). Some of these decisions were not present at all in the EEA, and we have had no time nor learning trials to acquire extensive personal frequency information about them. Instead, we need to make certain logical and probabilistic inferences using various rules of inference and, most importantly, we must decouple myriad sources of information that our autonomously functioning modules might be detecting and feeding into the decision (“No, the likeability of this salesperson should not be a factor in my deciding on this $25,000 car!”).

In fact, some of the System 1 heuristics that are in place might seriously subvert instrumental goals in a modern technological society. For example, one chapter in the Gigerenzer and Todd (1999) book (see Goldstein & Gigerenzer, 1999) is devoted to the so-called recognition heuristic—the chapter subheading being “How Ignorance Makes us Smart”. The idea behind such “ignorance-based decision making” as they call it, is the fact that some items of a subset are unknown can be exploited to aid decision making. The yes/no recognition response can be used as a frequency estimation cue. With ingenious simulations, Goldstein and Gigerenzer (1999) demonstrate how certain information environments can lead to such things as less-is-more effects: where those who know less about an environment can display more inferential accuracy in it.

One is certainly convinced after reading material like this that the recognition heuristic is certainly efficacious in some situations. But then one immediately begins to worry when we ponder how it relates to a market environment specifically designed to exploit it. If the first author of this chapter left his home—located in the middle of the financial and industrial capital of a developed country—and relied solely on the recognition heuristic, he could easily be led to:

(1) Buy a $3 coffee when in fact a $1.25 one would satisfy him perfectly.
(2) Eat in a single snack the number of grammes of fat he should have in an entire day.
(3) Pay the highest bank fees (because the highest fees are charged by the most recognized banks in Canada).

7. EVOLUT

(4) Incure credit card debt far
(5) Buy a mutual fund with

None of these behaviours serve none of them help get him to the recognition heuristic trigger while trying to make his way through

The proponents of ecological side of the ecological approach and Gigerenzer (1999) take an example of a group of mutual funds over a 6-month period as a pure artifact of an even large capitalization stocks out 2000). The adaptive heuristics repealed the basic laws of investing: longer time periods small capitalization large capitalization counterparty recognition for the large capital in that particular 6-month period risk/reward relationships to the

Borges et al. (1999) might known finding in the domain—of at 60% of consumers of fine consumer products and services are in the marketplace, whereas 60% of financial and consumer public

One leaves the writings of they take the term to mean—thologists, are being much too sure to achieve instrumental ratios goals of System 2 rather than t

THE UNACKNOWLEDGED

Why do the evolutionary psyche a tendency to misconstrue the interests of the replicators from is that the error follows fr
organism (see Over, 2000). With their usage, adopting what is being made. Ecological explanations of evolution (and the biological world) without accepting one implications (adapted for the EEA world) in the "one shot" affairs (job decisions, marriages) these decisions were not time nor learning trials to learn about them. Instead, we references using various rules (for example myriads of sources of knowledge "this person should be getting a k or a k") are in place might seriously systematic. For example, (his book (see Goldstein & Gigerenzer, 1999, on the phenomenon heuristic—"the chimpanzee is not a scientist"). The idea behind the human is that the fact that the environment can display more complex, more subtle, more artificial like this that the recognition cues. But then one immediately see that the author of this chapter left his and industrial capital of the recognition heuristic, he could satisfy him perfectly.

The fees of fat he should have in mind. The fees are charged by the

- Incur credit card debt rather than pay cash.
- Buy a mutual fund with a 6% sales charge rather than a no-load fund.

None of these behaviors serves his long-term instrumental goals at all—none of them help get him towards his reflectively acquired aspirations. Yet the recognition heuristic triggers these and dozens more that will trip him up while trying to make his way through the maze of modern society.

The proponents of ecological rationality refuse to acknowledge this downside of the ecological approach. For example, Borges, Goldstein, Ortmann, and Gigerenzer (1999) take curbs pride in the finding that a portfolio of stocks recognized by a group of Munich pedestrians beat two benchmark mutual funds over a 6-month period during the mid-1990s. This finding is of course a pure artifact of an extraordinary short period in the 1990s when large capitalization stocks outperformed small capitalization stocks (Over, 2000). The adaptive heuristics investigated by Borges et al. (1999) haven't repealed the basic laws of investing. Risk is still related to reward, and over longer time periods small capitalization stocks outperformed their less-risky large capitalization counterparts. Obviously, the Munich pedestrians had better recognition for the large companies—precisely those enjoying a good run in that particular 6-month period (which is, of course, too short for various risk/reward relationships to show themselves).

Borges et al. (1999) might alternatively have focused on another well-known finding in the domain of personal finance discussed by Bazerman (2001)—that consumers of financial services overwhelmingly purchase high-cost products that underperform in terms of investment return the low-cost strategies recommended by true experts (e.g., dollar-cost averaging into no-load index mutual funds). The reason is, of course, that the high-cost fee-based products and services are the ones with high immediate recognizability in the marketplace, whereas the low-cost strategies must be sought out in financial and consumer publications.

One leaves the writings of the ecological rationality theorists—whatever they take the term to mean—thinking that they, like the evolutionary psychologists, are being much too sanguine about the ability of System 1 processes to achieve instrumental rationality—to optimize the broad and reflective goals of System 2 rather than the short-leash evolutionary goals of System 1.

THE UNACKNOWLEDGED IMPORTANCE OF THE MEME

Why do the evolutionary psychologists and ecological theorists show such a tendency to misconstrue human rationality—to fail to distinguish the interests of the replicators from the interests of the vehicle? Our conjecture is that the error follows from a particular overgeneralization that is
encouraged by the evolutionary psychologists' attack on what Tooby and Cosmides (1992) call the Standard Social Science Model (SSSM). These evolutionary psychologists believe that the SSSM stands in the way of a fully articulated evolutionary psychology based on the assumption that the human brain is composed of content-specific information-processing mechanisms that have evolved as adaptations. Instead, the SSSM has become the default model of most social scientists in their view, and the SSSM contains many misleading assumptions about human cognition, including the assumption that the human mind is structured as a general, unbiased learning mechanism. Among the many assumptions of the SSSM that are wrong according to Tooby and Cosmides (1992), are its assumptions about culture. Tooby and Cosmides (1992) feel that the default assumptions of the SSSM of most social scientists are that "the individual is the more or less passive recipient of her culture and is the product of that culture" (p. 32) and that "human nature is an empty vessel, waiting to be filled by social processes" (p. 29).

The idea of free-floating cultural products—those totally unconditioned by and unadapted to evolved mental mechanisms (what Tooby & Cosmides, 1992 call epidemiological culture)—is an anathema to many evolutionary psychologists. In fact, Tooby and Cosmides (1992, p. 119) labour hard to convince the reader that even epidemiological culture "is also shaped by the details of our evolved psychological organization". To use Dawkins' (1976) term, evolutionary psychologists are hostile to the concept of the meme. But because human rationality is in large part a memetic product—a set of cultural tools for the Gregorian mind—evolutionary psychologists are prone to miss or denigrate its importance.

Dawkins (1976) introduced the term meme to refer to a unit of cultural information that is meant to be understood in rough (rather than one-to-one) analogy to a gene. Blackmore (1999) defines the meme as the instructions for behaviours and communications that can be learned by imitation broadly defined (in the sense of copying by the use of language, memory, or any other mechanism) and that can be stored in brains (or other storage devices). Collectively, genes contain the instructions for building the bodies that carry them. Collectively, memes build the culture that transmits them. Like the gene, the meme is a true replicator in the sense of the distinction made in theoretical biology between replicators and interactors (Dawkins, 1976; Hull, 1988; Sterelny, 2001; Sterelny & Griffiths, 1999; Williams, 1985, 1992). Replicators are entities that pass on their structure relatively intact after copying and interactors or vehicles are "those entities that interact as cohesive wholes with their environments in such a way that they can always be spellbound, or have greater fidelity, or have greater ion, future generations. Or, as Blackmore observes, 'genes want X' can always be spelled out. This is the only way they can spell out this sense that they are selfish'.

Memes are independent replicators that help the vehicle (those through having a higher fidelity)—replicators. The fundamental insight that the meme may play a role in the evolution of the vehicle that the human mind is not, which is, often use the example of the "host" so as to suggest the idea of a brain parasite, it may survive because of its good genes and the more the hosts, "what we have not previously died the way it has, simplicity, because psychologists say, that when you literally parasitize my brain, that might be the worst thing that a virus could do" (p. 192). With Dawkins' point in mind, writings of the memetic theorists 1993; Dennett, 1991, 1995; Lynott, 2001:

(1) Memes survive and spread that store them (most mem. would be in this category)
(2) Memes become frequent domain-specific evolution
attack on what Tooby and Cosmides (SSSM). These evo-
stands in the way of a fully n the assumption that the information-processing mech-
ism, the SSSM has become the view, and the SSSM contains n cognition, including the as a general, unbiased learn-
ning of the SSSM that are wrong is assumptions about culture. Its assumptions of the SSSM al is the more or less passive hat culture" (p. 32) and that be filled by social processes"
-those totally unconditioned ns (what Tooby & Cosmides, theme to many evolutionary 1992, p. 119) labour hard to culture “is also shaped by the on”. To use Dawkins’ (1976) he concept of the meme. But memetic product—a set of many psychologists are prone
to refer to a unit of cultural ough (rather than one-to-one) meme as the instructions for earned by imitation broadly iguage, memory, or any other r other storage devices). Col-
siding the bodies that carry at transmits them. Like the e of the distinction made in ractors (Dawkins, 1976; Hull, Williams, 1985, 1992). Repli-
atively intact after copying at interact as cohesive wholes
with their environments in such a way as to make replication differential” (Hull, 1988, p. 27).

The key idea in memetic theory is that the meme is a true selfish replicator in the same sense that a gene is—it acts only in its own “interests”. The anthropomorphic language about genes and memes having interests is shorthand for the complicated description of what is actually the case: that genes/memes that perform function X make more copies of themselves, copy with greater fidelity, or have greater longevity—and hence will leave more copies in future generations. Or, as Blackmore (1999, p. 5) states it, “the shorthand ‘genes want X’ can always be spelled out as ‘genes that do X are more likely to be passed on.’ This is the only power they have—replicator power. And it is in this sense that they are selfish”.

Memes are independent replicators. They do not necessarily exist in order to help the vehicle (those who hold the belief), they exist because through memetic evolution they have displayed the best fecundity, longevity, and copying fidelity—the defining characteristics of successful replicators. The fundamental insight triggered by memetic theory is that a meme may display fecundity and longevity without necessarily being true or helping the vehicle (the human being holding the belief) in any way. Memetic theorists often use the example of a chain letter. Here is a meme: “If you do not pass on this message to five people you will experience misfortune”. This is an instruction for a behaviour that can be copied and stored in brains. It survives because of its own self-replicating properties (it is neither good for the genes or for the vehicle). Dawkins (1976, p. 27) argues that “what we have not previously considered is that a cultural trait may have evolved in the way it has, simply because it is advantageous to itself”. Memetic theory asks instead what is it about certain memes that leads them to collect many “hosts” for themselves. Indeed, this type of language was suggested by Dawkins (1976, p. 192) himself who, paraphrasing Nick Humphrey, said that “when you plant a fertile meme in my mind you literally parasitize my brain, turning it into a vehicle for the meme’s propagation in just the way that a virus may parasitize the genetic mechanism of a host cell” (p. 192).

With Dawkins’ point in mind we are now in a position to extract from the writings of the memetic theorists (Aunger, 2000; Blackmore, 1999; Dawkins, 1993; Dennett, 1991, 1995; Lynch, 1996) a taxonomy of reasons for meme survival:

1. Memes survive and spread because they are helpful to the interactors that store them (most memes that reflect true information in the world would be in this category).

2. Memes become frequent because they fit genetic predispositions, or domain-specific evolutionary modules (this is the evoked culture that
is emphasized by evolutionary psychologists; see Cosmides & Tooby, 1992; Sperber, 1996).
(3) Memes spread because they facilitate the spread of the genes that make good hosts for these particular memes (religious beliefs that urge people to have more children would be in this category, see Lynch, 1996).
(4) Memes survive and spread because of the self-perpetuating properties of the memes themselves.

We must consider these categories in the context of the fact that many of the intentional-level goals that humans have are meme-installed: they are the products of our culture, rather than installed by the genes that built the vehicle. A schematic that helps to understand our conception of the intentional-level goal structure of System 1 and System 2 in terms of which replicator is a source of the goal is portrayed in Figure 7.6 (again, absolute areas are guesses—for illustrative purposes only). The goal structure of System 1 is dominated by gene-installed goals. These are the short-lease goals discussed earlier—nearly universal in the sense that they are shared by most humans and not the result of the environmental history of the organism. They are not flexible or generic goals, but instead are content specific, situation specific, and hard-wired to trigger (disgust and repulsion to noxious smells and substances, and fear responses to animals like snakes, would be examples; see Buss, 1999; Rozin, 1996; Rozin & Fallon, 1987).

System 2, with its more generic genetic goals shared by most humans (your conspecifics) and with more specific environmental experiences, we distinguish between mimetic viruses (as in the Dawkins q acquired memetic goals—and reflectively, with full awareness reflectively acquired goals are p Dawkins refers to. They may not like the vehicle-sacrificing gene vehicle merely to propagate their)

The figure also indicates the from becoming System 1 goals (ing). Through practice, memetic goal hierarchy of System 1. “Aspire to do just this—to have a without much thought. These parasites that, because they a difficult to dislodge.

Of course, meme-derived goals for the vehicle too. A reflective acquired because it served vehicle (the genes' interests)—can become a part of Fig. 7.5 that might have presented. Why is there a small that serve the vehicle's interests goals instanitated in System 1 not they were serving the inte Darwinian creature represents higher-level goal states of System 1 through pr acquired goal-states might be advantages to the vehicle (adv a contrary gene-installed goals—particular memes becoming ins create the area depicted in Fig. ' interests only. We might say 1 humans reflects the outcome of System 2. This is why the goals simply recapitulate the structure.

What evolutionary psychological conceptualization is the notion of
System 2, with its more general, flexible goals is more evenly balanced with genetic goals shared by most humans (e.g. rise in the dominance hierarchy of your conspecifics) and with meme-installed goals that are the result of the specific environmental experience (and culture) of the individual. In Fig. 7.6 we distinguish between mimetically acquired goals that are "caught" like viruses (as in the Dawkins quote above)—what we call nonreflectively acquired memetic goals—and memetic goals that an individual takes on reflectively, with full awareness of their effects on the organism. The nonreflectively acquired goals are perhaps the equivalent of the parasites that Dawkins refers to. They may not actually be good for the individual but, just like the vehicle-sacrificing genes discussed previously, these memes use the vehicle merely to propagate themselves.

The figure also indicates that meme-acquired goals need not be barred from becoming System 1 goals (automatic, autonomous, and rapidly triggering). Through practice, mimetically installed goals can become lodged in the goal hierarchy of System 1. "Branding" and other advertising gimmicks aspire to do just this—to have a logo for X trigger the "must have X" response without much thought. These then become especially pernicious memes—parasites that, because they are not part of the reflective mind, become difficult to dislodge.

Of course, meme-derived goals that become part of System 1 can be good for the vehicle too. A reflectively acquired meme—one that was reflectively acquired because it served vehicle ends (perhaps even vehicle ends that thwart the genes' interests)—can become part of System 1 as well. This fact explains a part of Fig. 7.5 that might have seemed perplexing when that figure was first presented. Why is there a small section of area in System 1 representing goals that serve the vehicle's interests only? One might have thought that all of the goals instantiated in System 1 would reflect the genes' interests whether or not they were serving the interests of the vehicle—rather like that of the Darwinian creature represented in Fig. 7.3. However, the possibility of the higher-level goal states of System 2 becoming installed in the more rigid and inflexible System 1 through practice opens up a new possibility. Reflectively acquired goal-states might be memes that were taken on for their unique advantages to the vehicle (advantages that might accrue because they trump contrary gene-installed goals—"don't flirt with your boss's wife"). Those particular memes becoming instantiated in System 1 through practice would create the area depicted in Fig. 7.5—System 1 goal-states serving the vehicle's interests only. We might say that in situations such as this, System 1 in humans reflects the outcome of residing in a brain along with a reflective System 2. This is why the goal structure of System 1 in humans does not simply recapitulate the structure of a Darwinian creature depicted in Fig. 7.3.

What evolutionary psychologists do not like about the previous conceptualization is the notion of memes becoming completely "unglued" from
genetic control. Instead, evolutionary psychologists prefer Lumsden and Wilson’s (1981) notion that the genes hold culture on a leash (see Sperber, 1996). What they do not like is the idea, which we are advancing here, that at a certain level of recursiveness a Gregorian mind populated with cultural tools in the form of memeplexes designed for the evaluation of other memeplexes (science; logic; some notions from decision science such as consistency, transitivity, etc.) acquires some autonomy from genetic control. But it is just such autonomy that we are arguing for. We are in fact arguing that the cultural tools of logic and decision science, when reflectively used in conjunction with the potent cultural insight that there can be a conflict of interest between replicators and vehicle, have the potential to create a creature with a uniquely critical and discerning type of self reflection. In short, we are arguing that understanding the full implications of the replicator/vehicle distinction may be a cultural tool that could foster even greater levels of self reflection than humans have heretofore achieved.

Combined with the tools of decision science, the vehicle/replicator distinction can spawn thoughts and new tools for the restructuring of human goals—new memes that further sever the connection between memeplexes resident in some brains and genetic goals installed by the replicators. Indeed, we propose that this cultural change is already underway. There are already memeplexes in the air (of which this book is one) that will, contrary to the emphasis in the writings of evolutionary psychologists, further background the role of the genes in human culture.6

Evolutionary psychologists resist this extrapolation, falling back on their “culture on a leash” notion. For example, Tooby and Cosmides (1992, p. 119) insist that “epidemiological culture is also shaped by the details of our evolved psychological organization” and, even more strongly, that “our developmental and psychological programs evolved to invite the social and cultural worlds in, but only the parts that tended, on balance, to have adaptively useful effects” (p. 87). But the evolutionary psychologists seem to have underestimated the power of the memes to break this linkage. What they have neglected is the recursive power of evaluative memes in the context of an organism that has become aware of the replicator/vehicle distinction. Science writer Robert Wright (1994) paraphrases the Tooby and Cosmides statement above about “our developmental and psychological programs evolved . . .”

6 Indeed, genetic engineering for purposes of human health and longevity is perhaps the ultimate triumph of Dawkins’ (1976) so-called “survival machines” (the human vehicles) over their creators—the replicators. With the technology of genetic engineering, we, who were built by the replicators to serve as their survival machines, use them for our own goals—goals that are not the genes’ goals (e.g. survival past our reproductive years). Williams (1988) uses such an example to counter Stent’s (1978) argument against Dawkins (1976) that rebellion against one’s own genes is a contradiction. Williams (1988, p. 403) notes that Stent “apparently missed the relevance of major technologies (hair dyeing, tonsillectomy, etc.) based on such rebellion”.

into the more readable notion that the brains they settle into” (Wrig Cosmides (1992), Wright (1994) low from becoming aware of the distinction. So, after noting that it brains they settle into, Wright (they’re good for those brains in it)

A brain that realizes this sta begin the (admittedly difficult) t genes when the culture is dysfun replicator/vehicle distinction an such as science, logic, and decis vehicle-thwarting goals from i mematic structures that serve course, this is exactly what rationality—largely a product accomplish.

This might seem like a Prior cognitive science has emphasisi tific knowledge result in chang land (1989, 1995) has long en change our folk language of’ t have emphasized how moral t deterministic explanatory pows (Wright, 1994). Already, among violations of transitivity and be a cause of cognitive sans unprecedented. A full appre vehicle distinction, with its emp to the personal and subpers versus genetic fitness), could be. Of course, we do not mean t guilty of committing a sophist current function from ancestr example, Pinker (1997, p. 401) view and explicitly recognizes t replicators and the vehicle: “G recipes for making the brain ar live in a parallel universe, scar
logists prefer Lumsden and Fisher on a leash (see Sperber, we are advancing here, that at nimbly populated with cultural knowledge of other meme-plexes such as consistency, genetic control. But it is just in fact arguing that the cycle-effectively used in conjunction with a uniquely short, we are arguing that replicator/vehicle distinction may need levels of self reflection that

the vehicle/replicator distinction the restructuring of human notion written between memeplexes led by the replicators. Indeed, underway. There are already one that will, contrary to the sologists, further background

6 solution, falling back on their y and Cosmides (1992, p. 119) haped by the details of our en more strongly, that "our olved to invite the social and ed, on balance, to have adap-

ty psychologists seem to have k this linkage. What they have memes in the context of an or/vehicle distinction. Science 60y and Cosmides statement ogical programs evolved..."

health and longevity is perhaps the machines" (the human vehicles) over tic engineering. we, who were built by for our own goals—goals that are not Williams (1988) uses such an example 76 that rebelling against one's own that Senn "apparently missed the ; etc.) based on such rebellion".

into the more readable notion that ideas must "have a kind of harmony with the brains they settle into" (Wright, 1994, p. 366). However, unlike Tooby and Cosmides (1992), Wright (1994) realizes that there are implications that follow from becoming aware of this fact in the context of replicator/vehicle distinc-
tion. So, after noting that ideas must have a kind of harmony with the brains they settle into, Wright (1994, p. 366) warns that "that doesn't mean they're good for those brains in the long run".

A brain that realizes this startling (and still underappreciated) fact might begin the (admittedly difficult) process of shipping culture off the leash of the genes when the culture is dysfunctional for the person.6 A brain aware of the replicator/vehicle distinction and in the possession of evaluative memeplexes such as science, logic, and decision theory might begin a process of pruning vehicle-thwarting goals from intentional-level psychology and reinstalling memetic structures that serve the vehicle's interests more efficiently (of course, this is exactly what the canons of normative instrumental rationality—largely a product of the twentieth century—were designed to accomplish).

This might seem like a Promethean goal but, in fact, a rich tradition in cognitive science has emphasized how cultural changes and increased scientific knowledge result in changes in folk psychology. For example, Churchill (1989, 1995) has long emphasized how a mature neuroscience might change our folk language of the mental and of behaviour. Other theorists have emphasized how moral notions change as general knowledge of the deterministic explanatory power of neuroscience becomes more widespread (Wright, 1994). Already, among educated citizens of the twenty-first century, violations of transitivity and independence of irrelevant alternatives can be a cause of cognitive sanction in ways that are probably historically unprecedented. A full appreciation of the implications of the replicator/vehicle distinction, with its emphasis that differing optimization criteria apply to the personal and subpersonal levels of analysis (utility maximization versus genetic fitness), could have equally profound cultural implications.

Of course, we do not mean to imply that all evolutionary psychologists are guilty of committing a sophisticated version of the genetic fallacy (inferring current function from ancestral function, see Dennett, 1995, p. 465). For example, Pinker (1997, p. 401) does not endorse the culture-on-a-short-leash view and explicitly recognizes the implications of the differing interests of the replicators and the vehicle: "Genes are not puppetmasters; they acted as the recipes for making the brain and body and then they got out of the way. They live in a parallel universe, scattered among bodies, with their own agendas."

---

6 Those not committed a priori to a relativistic denial of the notion of cultural advance might well argue that the history of civilization reflects just this trend (the emancipation of women and the control of our reproductive lives come immediately to mind).
Clearly, not all evolutionary psychologists miss the implication of replicator/vehicle distinction for conceptions of rationality. But some evolutionary theorists do—quite egregiously. In an astonishing essay titled "How Evolutionary Biology Challenges the Classical Theory of Rational Choice", Cooper (1989) basically argues that when choosing between your own goals and those of your genes, you should opt for the latter! After a marvellous discussion of why a probability-matching strategy (Estes, 1964, 1976) might be a fitness-optimizing rather than the utility-maximizing strategy (picking the most frequent option each time), Cooper (1989, p. 459) implies that this outcome undermines the prescriptive force of the utility maximizing strategy: "The upshot is that one is faced with a dilemma. Either rationality is not always the fittest policy, or else classical decision analysis is not as universally rational as is commonly claimed. If the latter horn of the dilemma is seized (and I shall argue that that is indeed the lesser of the evils) . . . .". Of course, early in the article one feels that this is a verbal slip. But, ten pages on, we find out that the author does indeed want to argue that we should follow goals that satisfy our genes rather than as individuals as individual organisms. The ordinary application of the logic of decision science is termed "naively applied" when interpreted "with the individual treated as an isolated locus of decision making and with the role of the genotype ignored" (Cooper, 1989, p. 473). The instability in preference orderings that signal the failure of individual utility maximization (Dawes, 1998; Kahneman & Tversky, 2000; Slovic, 1995) are defended because "perhaps some of the observed instability is due to adaptive strategy mixing. If so, instability would have to be reevaluated; when one is acting as an agent of one's genotype, it could sometimes be a sound strategy" (Cooper, 1989, p. 473). But who in the world would want to act as an agent of one's genotype rather than in the service of one's own life goals! This is precisely the choice Cooper (1989) is positing when he puts the concerns of genetic fitness against those of instrumental rationality.

Let the reader worry that we caricature, Cooper (1989) leaves us in no doubt because he concludes his paper with replies to possible criticisms of his view. In this section he makes it clear that his view is that "all adequate choice rules are seen as mere extensions of evolutionary principles" (p. 475) because "the maximization of fitness is a ubiquitous goal" (p. 475). A ubiquitous goal of the genes no doubt, but—completely ignoring the individual organism as a potential critical locus of utility maximization—Cooper (1989, p. 475) proposes that we "revise the classical theory [utility theory] itself at its mathematical core, letting biologically motivated decision rules . . . replace or supplement the traditional ones as basic decision rules"

In his summary statement, Cooper (1989, p. 479) makes it clear that the proposition he wishes to defend is that "the traditional theory of rationality is invalid as it stands, and in need of biological repair", and acknowledges that this is "a stance not likely to please theorists, but perhaps underlines the need for others that have the power of the modern evolutionary champions the notion that fitness or "biologically naive" (p. 480). L (1989) seems to have taken the view that evolutionary psychologists so explicitly throw off evolutionary psychologists at (Cosmides & Tooby, 1994a; Gi actually do it implicitly in the contention that "the tradition of and in need of biological repair' and evolutionary psychologismic Cooper's (1989) view w sions suggest that traditiona rationality need to be reexam Throughout this essay they re fround claims to the contrary, t e.g. because of processing u rational" (p. 329).

It is, in fact, relatively con rational thought to be desegni biases approach. Gigerenzer a (1989) extreme position in the traditional rationality as a universal i of 'good' reasoning on which t biases views were built" (p. 65 so much useless "baggage" in the abilities of the mind and us to see how thought process and probability can solve rea

Likewise, as noted in our c that cognitive illusions (viola "disappear" with more evol imply that these normative v they (the normative violation does not follow them does not what these authors want us to only rationality that need
that this is "a stance not likely to be popular with confirmed classical decision theorists, but perhaps understandable to evolutionists, psychologists, philosophers, and others that have been impressed by the pervasive explanatory power of the modern evolutionary perspective" (p. 479). The view explicitly championed is the notion that "behavioral rationality [be] interpreted in terms of fitness" (p. 480) and that any dissent from this policy be viewed as "biologically naive" (p. 480). Like the sociobiologists before him, Cooper (1989) seems to have taken the defense of the genes as his brief!

Cooper's (1989) view may well seem extreme, and few evolutionary psychologists so explicitly throw out the vehicle with the bathwater. But many evolutionary psychologists and proponents of ecological rationality (Cosmides & Tooby, 1994a; Gigerenzer, 1996a; Gigerenzer & Todd, 1999) actually do do it implicitly in the way that they echo Cooper's (1989, p. 479) contention that "the traditional theory of rationality is invalid as it stands, and in need of biological repair". For example, in a paper discussing economics and evolutionary psychology, Cosmides and Tooby (1994a) quite closely mimic Cooper's (1989) view when they argue that "evolutionary considerations suggest that traditional normative and descriptive approaches to rationality need to be reexamined" (Cosmides & Tooby, 1994a, p. 329).

Throughout this essay they repeat the odd declaration that "despite widespread claims to the contrary, the human mind is not worse than rational (e.g. because of processing constraints)—but may often be better than rational" (p. 329).

It is, in fact, relatively common for the traditional normative rules of rational thought to be denigrated in the literature critical of the heuristics and biases approach. Gigerenzer and Goldstein (1996) adopt exactly Cooper's (1989) extreme position in their argument that in their view "questions classical rationality as a universal norm and thereby questions the very definition of 'good' reasoning on which both the Enlightenment and the heuristics-and-biases views were built" (p. 651). The classical norms are referred to as just so much useless "baggage" in quotes such as the following: "A bit of trust in the abilities of the mind and the rich structure of the environment may help us to see how thought processes that forgo the baggage of the laws of logic and probability can solve real-world adaptive problems quickly and well" (Gigerenzer & Todd, 1999, p. 365).

Likewise, as noted in our discussion above, Gigerenzer's repeated refrain that cognitive illusions (violations of the canons of normative rationality) "disappear" with more evolutionarily propitious problem representations implies that these normative violations are of no concern. Since we know that they (the normative violations) are of concern to the vehicle (a vehicle who does not follow them does not maximize utility), we can only conclude that what these authors want us to imply is that evolutionary rationality is the only rationality that need concern us—precisely Cooper's (1989) point.
(although the point is much more subtle and somewhat hidden in the writings of the ecological theorists).

Cosmides and Tooby (1994a), in an essay directed at economists, ignore completely the role of memetic evolution and culture in determining human preferences. In a series of points laid out like a series of axioms they argue that because "natural selection built the decision-making machinery in human minds" (p. 328) and because "this set of cognitive devices generates all economic behavior", "therefore . . . the design features of these devices define and constitute the human universal principles that guide economic decision making" (p. 328).

These postulates lead Cosmides and Tooby (1994a, p. 311) to the grandiose claim that "evolutionary psychology should be able to supply a list of human universal preferences, and of the procedures by which additional preferences are acquired or reordered". But to the extent that the claim is true, it is only because the grain-size of the predictions will be all wrong. The economic literature is not full of studies debating whether humans who are dying of thirst prefer water or shelter—or whether men prefer 23-year-old females over 75-year-old ones. Instead, the literature is full of studies trying to determine the rationale for such fine-grained judgements as, for example, whether a poor briefcase produced by an athletic shoe company will adversely affect the family brand name (Abuwaqa & Gurhan-Canli, 2000). Economists and psychologists are not debating the reasons for preferences among basic biological needs. Instead, they are debating the reasons for fine-grained preferences among highly symbolic products embedded in a complex, information-saturated, "attention-based" (Davenport & Beck, 2001) economy. Even after we grant evolutionary assumptions like, for example, that people use clothes purchases for some type of modern dominance display or sexual display, we have not progressed very far in explaining how brand names wax and wane in the fashion world, or how price elastic such purchases will be, and/or what kind of substitutability there will be among these types of goods.

This essay by Cosmides and Tooby (1994a) directed to economists serves to reinforce all of the worst Panglossian tendencies in the latter discipline. For example, Kahneman, Wakker, and Sarin (1997) discuss why experienced utility is essentially ignored in modern economics despite psychological studies showing that experienced utility is not identical to expected utility. They argue that experienced utility is ignored by economists on the grounds that "choices provide all necessary information about the utility of outcomes because rational agents who wish to do so will optimize their hedonic experience" (Kahneman et al., 1997, p. 375). Two-process theories of cognition—in conjunction with the assumptions that we have made about goal structures—help to explain why this assumption might not hold. The choices triggered by the goal structures of System 1 might not always be oriented towards the optimization of hedonic experience it just a means to an end (largely genetic goals). This 5th pleasurable if ultimate fitness goal

CHOOSING THE REPPLICATORS: EVOLUTION WITHOUT G

What the evolutionary psych have occasionally been guilty of and (1995, p. 82) has termed "the (1995), in their "zeal to exploit the tendency to "skip whole layers of thing securely and neatly to to appeal to reductionist eft impressed with the seminal a Table 1 of Buss, Haselton, S long list of important behav uncovered because of apply psychology and consider its during the 1990s (Barlow, C Cartwright, 2000: Cosmides Pinker, 1997; Plotkin, 1998) t of rationality, the evolutions They too easily gloss over t mismatches and their implic art intelligence and/or gener ous System 1 responses (g Because many of the tools are cultural inventions (meme in technological societies psychologis.

In the extreme, evolution siding with genetic interests: essay quoted extensively abo betting against the probabilit fare" (p. 477), but argues th individual identifies its own well what if? Then yes, maybe the people with such loyalty to t type? Which alleles, for exat ings for? Beyond a few scie
somewhat hidden in the writings directed at economists, ignore culture in determining human a series of axioms they argue cision-making machinery in cognitive devices generates all features of these devices define that guide economic decision / (1994a, p. 311) to the gran- ought be able to supply a list occlusions by which additional the extent that the claim is fictions will be all wrong. The ing whether humans who are ether men prefer 23-year-old nature is full of studies trying l judgements as, for example, t shoe company will adversely Jurhan-Cani, 2000). Econo- asons for preferences among g the reasons for fine-grained s embedded in a complex, enport & Beck, 2001) eon- tions like, for example, that modern dominance display or ar in explaining how brand how price elastic such pur- ify there will be among these directed to economists serves es in the latter discipline. For discuss why experienced util- despite psychological studies al to expected utility. They nomists on the grounds that out the utility of outcomes ptimize their hedonic experi- essories of cognition—in ade about goal structures— ed. The choices triggered by ies be oriented towards the optimization of hedonic experience for the individual agent. The hedonic experience is just a means to an end for most of the goals lodged in System 1 (largely genetic goals). This System will readily sacrifice the vehicle’s hedonic pleasure if ultimate fitness goals are achievable without it.

CHOOSING THE VEHICLE RATHER THAN THE REPLICATORS: EVOLUTIONARY PSYCHOLOGY WITHOUT GREEDY REDUCTIONISM

What the evolutionary psychologists and ecological rationality theorists have occasionally been guilty of in the domain of rationality is what Dennett (1995, p. 82) has termed “greedy reductionism”. According to Dennett (1995), in their “zeal to explain too much too fast”, greedy reductionists tend to “skip whole layers or levels of theory in their rush to fasten everything securely and neatly to the foundation” (p. 82). Like Dennett (1995), we applaud reductionist efforts in the behavioural sciences. We are impressed with the seminal achievements of evolutionary psychology (see Table 1 of Buss, Haselton, Shackelford, Belske, & Wakefield, 1998, for a long list of important behavioural relationships that were in large part uncovered because of applications of the theoretical lens of evolutionary psychology) and consider its emergence as a dominant force in psychology during the 1990s (Barkow, Cosmides, & Tooby, 1992; Buss, 1999, 2000; Cartwright, 2000; Cosmides & Tooby, 1994b; Geary & Bjorklund, 2000; Pinker, 1997; Plotkin, 1998) to be a salutary development. But in the area of rationality, the evolutionary psychologists have built a bridge too far. They too easily gloss over the important issue of replicator/vehicle goal mismatches and their implications. They too easily dismiss the role of general intelligence and/or general computational power in overriding deleterious System 1 responses (Stanovich, 1999; Stanovich & West, 2000). Because many of the tools of instrumental and epistemic rationality are cultural inventions (memes) and not biological modules, their usefulness in technological societies is too readily dismissed by evolutionary psychologists.

In the extreme, evolutionary theorists begin to sound as if they are siding with genetic interests against those of people. Cooper (1989), in the essay quoted extensively above, admits that “nonclassical behaviors such as betting against the probabilities are detrimental to the reasoner’s own welfare” (p. 477), but argues that this is justified because “what if the individual identifies its own welfare with that of its genotype?” (p. 477). Well, what if? Then yes, maybe they should probability match. But who are these people with such loyalty to the random shuffle of genes that is their geno- type? Which alleles, for example, do you have particularly emotional feel- ings for? Beyond a few scientists too narrowly focused on the promised
explanatory power of evolutionary psychology, we doubt that there are such people.\footnote{To be precise, we are doubting whether there are people who say they value their genome and have an accurate view of what they are valuing when they say this. For example, in such a case, the person would have to be absolutely clear that valuing your own genome is not some proxy for valuing your children, be clear that having children does not replicate one’s genome; and be clear about the fact that the genome is a subpersonal entity.}

Gibbard (1990, pp. 28–29) offers the more reasoned view:

It is crucial to distinguish human goals from the Darwinian surrogate of purpose in the “design” of human beings... The Darwinian evolutionary surrogate for divine purpose is now seen to be the reproduction of one’s genes. That has not, as far as I know, been anyone’s goal, but the biological world looks as if someone quite resourceful had designed each living thing for that purpose... A person’s evolutionary telos explains his having the propensities in virtue of which he develops the goals he does, but his goals are distinct from this surrogate purpose. My evolutionary telos, the reproduction of my genes, has no straightforward bearing on what it makes sense for me to want or act to attain... A like conclusion would hold if I knew that I was created by a deity for some purpose of his: his goal need not be mine... Likewise, if I know that my evolutionary telos is to reproduce my genes, that in itself gives me no reason for wanting many descendants.

In short, “human moral propensities were shaped by something it would be foolish to value in itself, namely multiplying one’s own genes” (p. 327)

Gibbard’s (1990) view is shared by distinguished biologist George Williams, who feels that “there is no conceivable justification for any personal concern with the interests (long-term average proliferation) of the genes we received in the lottery of meiosis and fertilization. As Huxley was the first to recognize, there is every reason to rebel against any tendency to serve such interest” (Williams, 1988, p. 403).

Dennett (1995) discusses this point in a different way—by making the astonishing observation that, until quite recently, the genes were the only beneficiary of all of the selective forces on the planet. That is, “there were no forces whose principal beneficiary was anything else. There were accident and catastrophes (lightning bolts and tidal waves), but no steady forces acting systematically to favor anything but genes” (Dennett, 1995, p. 326). But now we are here. There exist in the universe, for the first time, another set of interests because, unlike Darwinian creatures, our interests are not necessarily those of our genes. Rationality is the meme that trumps genetic interests in cases such as this. The remarkable cultural project to advance human rationality concerns how to best advance human interests whether or not they coincide with genetic interests. Its emancipatory potential is lost if we fail to

\footnote{To be precise, we are doubting whether there are people who say they value their genome and have an accurate view of what they are valuing when they say this. For example, in such a case, the person would have to be absolutely clear that valuing your own genome is not some proxy for valuing your children, be clear that having children does not replicate one’s genome; and be clear about the fact that the genome is a subpersonal entity.}
7. EVOLUTIONARY VERSUS INSTRUMENTAL GOALS

see the critical divergence of interests that creates the distinction between evolutionary and instrumental rationality.

ACKNOWLEDGEMENTS

Preparation of this chapter was supported by a grant from the Social Sciences and Humanities Research Council of Canada to Keith E. Stanovich.

REFERENCES


Borges, B., Goldstein, D.G., Ottmann, A., & Gigerenzer, G. (1999). Can insurance beat the...
stock market? In G. Gigerenzer & P.M. Todd (Eds.), Simple heuristics that make us smart (pp. 59-72). New York: Oxford University Press.


7. EVOLUTIONARY VERSUS INSTRUMENTAL GOALS


Todt, P.M., Fidler, L., & Krauss, and Reasoning, 6, 375-384.

Todt, P.M., & Gigerenzer, G. (2000) and Brain Science, 23, 727-780.


Author Index

Abbey, A. 22
ABC Research Group, The 4, 88, 121
Abell, F. 77
Adler, J.E. 174–175, 195
Adolphs, R. 75–76, 87, 89, 91
Aggleton, J.P. 76
Ahlawalia, R. 218
Ahn, W. 165
Aiello, L.C. 77
Ajzen, I. 160, 184
Alexander, M.P. 75–76
Allan, L.G. 177
Allen, J. 116
Allison, T. 75, 77
Allman, J. 76
Altmor, A. 6, 35, 49, 50–52, 111–112, 126
Amaral, D.G. 76–77
Anderson, A. 76
Anderson, J.R. 171, 177–178, 180, 185, 199